# BAYESOVE MREŽE

## Uvoz podataka u GeNiE

Podaci u program GeNiE uveženi su iz datoteke „data\_bayes.csv“.

## Diskretizacija numeričkih varijabli

Pretvaramo sve numeričke (kontinuirane) varijable u kategoričke varijable.

* Varijabla “time”

A screenshot of a computer

Description automatically generated

Slika 1. Diskretizacija varijable “time”.

* Varijabla “age”

A screenshot of a computer

Description automatically generated

Slika 2. Diskretizacija varijable "age".

* Varijabla “wtkg”

A screenshot of a graph

Description automatically generated

Slika 3. Diskretizacija varijable “wtkg”

* Varijabla “karnof”

A screenshot of a computer

Description automatically generated

Slika 4.Diskretizacija varijable "karnof"

* Varijabla “preanti”

A screenshot of a computer

Description automatically generated

Slika 5.Diskretizacija varijable "preanti"

* Varijabla “cd40”

A screenshot of a computer

Description automatically generated

Slika 6.Diskretizacija varijable "cd40".

* Varijabla “cd420”

A screenshot of a computer

Description automatically generated

Slika 7.Diskretizacija varijable "cd420".

* Varijabla “cd80”

A screenshot of a computer

Description automatically generated

Slika 8.Diskretizacija varijable "cd80".

* Varijabla “cd820”

A screenshot of a computer

Description automatically generated

Slika 9.Diskretizacija varijable "cd820".

## Kreiranje i treniranje modela

Pošto varijable “time”, “cd420” i “cd40” imaju najjači IV, stavit ćemo im direktnu vezu na ciljnu varijablu “cid”.

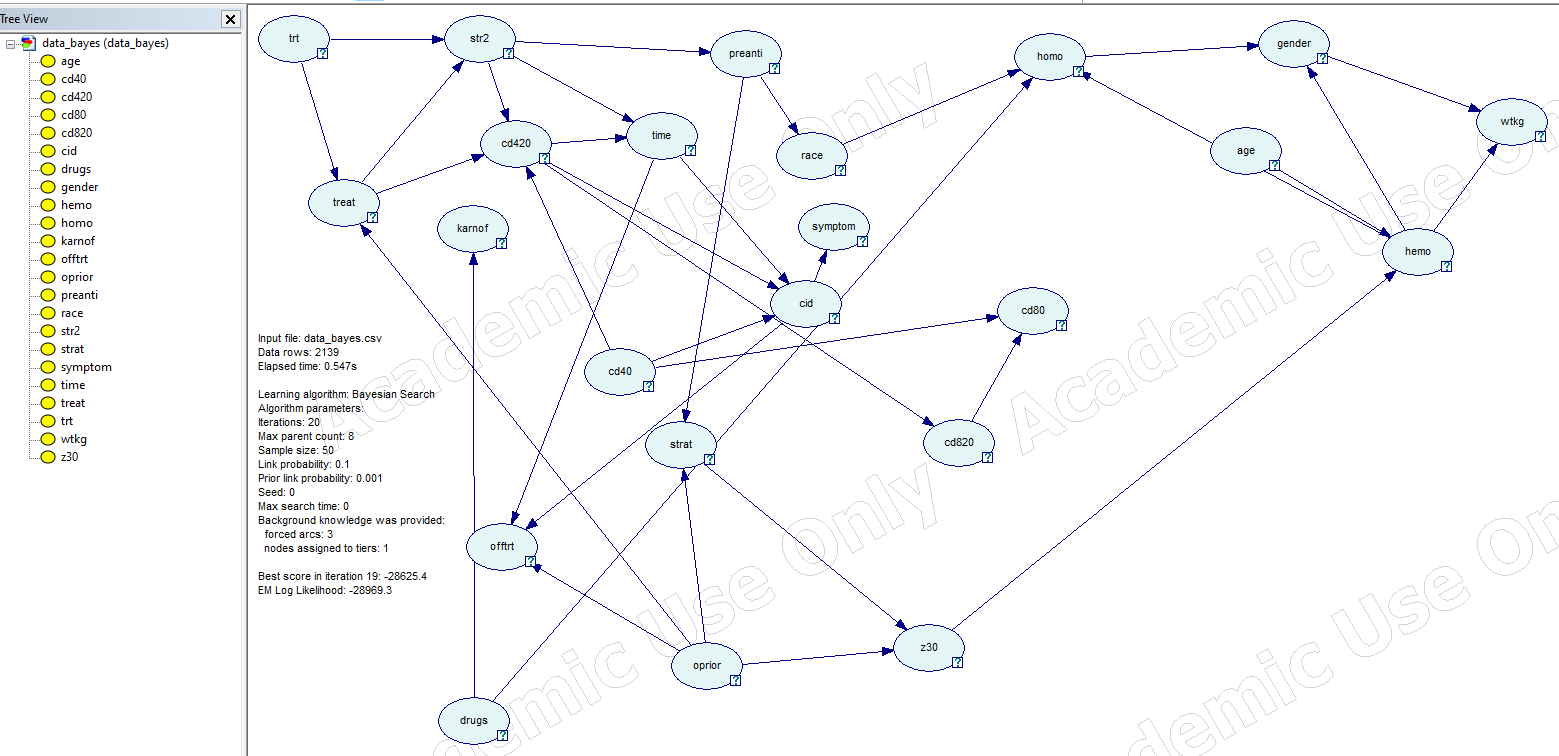
A screenshot of a computer

Description automatically generated

Slika 10. Postavljanje direktne veze prediktora sa jakim IV na ciljnu varijablu.

Uz postojeće veze nakon treniranja modela stvoreno je 32 novih veza:

1. varijable trt i str2
2. varijable trt I treat
3. varijable treat I str2
4. varijable str2 I time
5. varijable treat I cd420
6. varijable str2 I cd420
7. varijable str2 I preanti
8. varijable cd420 I time
9. varijable oprior I treat
10. varijable oprior I offtrt
11. varijable time I offtrt
12. varijable cd420 I cd820
13. varijable cd40 I cd420
14. varijable cid I offtrt
15. varijable cid I symptom
16. varijable preanti I race
17. varijable preanti I strat
18. varijable race I homo
19. varijable homo I gender
20. varijable drugs I homo
21. varijable cd40 I cd80
22. varijable cd820 I cd80
23. varijable drugs I karnof
24. varijable oprior I strat
25. varijable oprior I z30
26. varijable strat I z30
27. varijable z30 I hemo
28. varijable hemo I homo
29. varijable hemo I gender
30. varijable gender I wtkg
31. varijable hemo I wtkg
32. varijable age I hemo



Slika 11. Stvorene su nove veze nakon treniranja modela.

**cd40** - **cd420**:

cd40, koji predstavlja broj CD4 stanica na početku studije, izravno utječe na cd420, broj CD4 stanica nakon 20+/-5 tjedana. Ovo sugerira da početni broj CD4 stanica može biti prediktivni faktor za njihov broj nakon 20 tjedana liječenja.

**cd420** - **cd820/time**: Promjene u broju CD4 stanica (cd420) su povezane s promjenama u broju CD8 stanica (cd820) i vremenom do neuspjeha (time), što ukazuje na važnost praćenja oba seta imunoloških markera.

**cd420** **– cid**, **time – cid, cd40 -cid**:

Varijable **cd420**, **cd40** i **time** direktno su povezane s **cid**. Ovo upućuje na to da su kako broj početnih CD4 stanica, te broj CD4 stanica nakon 20 tjedana i vrijeme do neuspjeha ili cenzuriranja, značajni za ishod ciljne varijable.

**trt** - **str2/treat**: Varijabla **trt** (tretman) direktno je povezana s varijablom **str2** (antiretrovirusna povijest) i **treat** (indikator tretmana), što ukazuje da povijest liječenja i trenutni tretman imaju povezan utjecaj.

**treat** - **str2/cd420**: Varijabla **treat** je povezana sa **str2** i **cd420** (broj CD4 stanica nakon 20 tjedana), sugerirajući da trenutni tretman utječe na povijest liječenja i stanje imunološkog sustava.

**str2** (pacijent ranije bio izložen antiretrovirusnom liječenju.) - **time/cd420/preanti**: Varijabla **str2** je povezana s vremenom do neuspjeha (time), brojem CD4 stanica nakon 20 tjedana (cd420) i brojem dana antiretrovirusne terapije prije početka studije (preanti), implicirajući da prethodna izloženost antiretrovirusnom liječenju može utjecati na trajanje do neuspjeha, kao i na imunološki odgovor.

**oprior** - **treat/offtrt:** Prethodna ne-ZDV antiretrovirusna terapija (**oprior**) povezana je s trenutnim tretmanom i statusom prekida tretmana (**offtrt**), sugerirajući da ranije tretmane treba uzeti u obzir kada se razmatra trenutni ili budući tretman.

**cid** - **offtrt/symptom**: Ciljna varijabla **cid** (indikator neuspjeha/cenzuriranja) povezana je s prestankom tretmana (**offtrt**) i simptomatskim indikatorom (**symptom**), implicirajući da ti faktori igraju značajnu ulogu u određivanju ishoda pacijenta.

**race - homo/preanti**: Rasa (**race**) je povezana s homoseksualnom aktivnošću (**homo**) i prethodnom antiretrovirusnom terapijom (**preanti**), što može sugerirati postojanje sociodemografskih uzoraka.

Na temelju ove mreže, može se pretpostaviti da su početni i naknadni brojevi CD4 stanica (označeni varijablama **cd40** i **cd420**) i vrijeme (**time**) presudni za razumijevanje i predviđanje ishoda (**cid**). Primjerice, pacijent s višim brojem CD4 stanica na početku studije i nakon 20 tjedana liječenja može imati bolje dugoročne ishode u smislu izbjegavanja "neuspjeha" ili dužeg perioda prije cenzuriranja.

## Prikaz jačina veza/utjecaja

A network of information on a white background

Description automatically generated with medium confidence

Slika 12. Prikaz jačine utjecaja.

Varijabla **str2** (antiretrovirusna povijest) s varijablom **preanti** (dani antiretrovirusne terapije prije studije):

* Najjača povezanost između ove dvije varijable pokazuje da je povijest prethodnog izlaganja antiretrovirusnom liječenju direktno povezana s brojem dana na koje su pacijenti bili tretirani prije sudjelovanja u studiji.

Varijabla **preanti** i **strat**:

* Snažna povezanost između trajanja prethodne antiretrovirusne terapije (**preanti**) i kategorizacije prema dužini prethodnog liječenja (**strat**) sugerira da kako pacijenti dulje uzimaju terapiju, tako se mijenja i njihova kategorizacija u kontekstu studije.

Varijabla **hemo** (Hemofilija) na varijablu **homo** (Homoseksualna aktivnost) i varijabla **oprior** (Prethodna ne-ZDV antiretrovirusna terapija) na varijablu **z30** (ZDV upotreba u 30 dana prije početka studije):

* Veza između varijable **hemo** (hemofilija) i varijable **homo** (homoseksualna aktivnost) može ukazivati na to da su ljudi s hemofilijom imali aktivane homoseksualne aktivnosti. Što se tiče veze između varijable oprior (prethodna ne-ZDV terapija) i **z30** (upotreba ZDV-a u 30 dana prije početka studije), izgleda da ako su ljudi ranije koristili druge lijekove osim ZDV-a, postoji veća šansa da će koristiti ZDV u trenutku kad počinje studija.

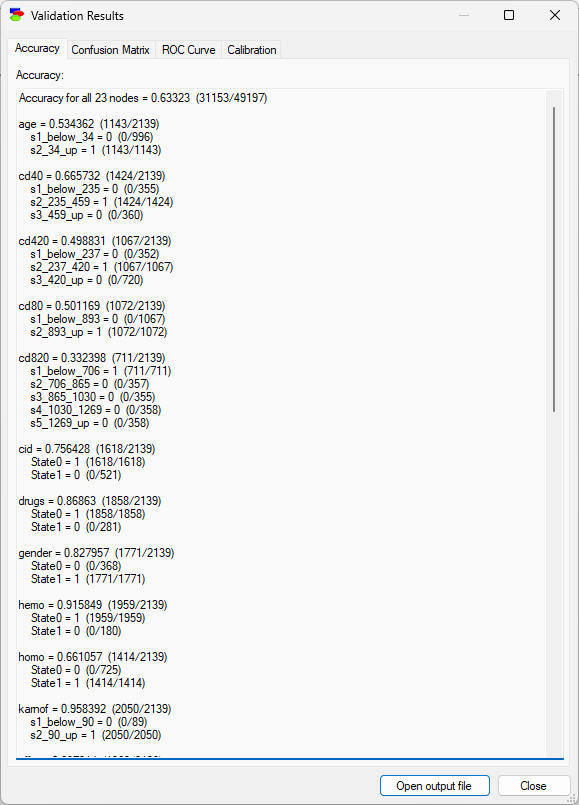
Varijabla **cd40** (CD4 broj na početku studije) na varijablu **cd420** (CD4 broj na 20±5 tjedana), varijabla **cid** (Indikator cenzuriranja) na **offtrt** (prestanak tretmana) i varijabla **time** (Vrijeme do neuspjeha ili cenzuriranja) na **offtrt**:

* Ako su prve vrijednosti CD4 visoke, to može značiti da će i kasnije biti visoke. Ako se podaci o pacijentu prestanu pratiti (cid), to može utjecati na to hoće li pacijent nastaviti s liječenjem. Također, ako pacijent brzo postane lošije ili mu liječenje ne pomaže (**time**), može se dogoditi da prije prekine s liječenjem.

Varijabla **time** na varijablu **cid**:

* ova veza pokazuje da dulje vremensko razdoblje prije neuspjeha ili cenzuriranja izravno utječe na ishod liječenja. Dugotrajnost do neuspjeha obično je povezana s uspješnim odgovorom na liječenje, dok kraće vremensko razdoblje može upućivati na to da tretman nije bio efikasan.

## „Skoriranje“ rezultata, ROC krivulja i matrica zbunjenosti



Slika 13. Matrica zbunjenosti.

Matrica zbunjenosti prikazuje broj točno i pogrešno klasificiranih zapisa. Ovo govori koliko nam model dobro radi.

Primjer za varijablu “time”.

A screenshot of a computer

Description automatically generated

Slika 14. Matrica zbunjenosti za varijablu "time".

Primjer ROC krivulje za varijablu “time”. Površina ispod ROC krivulje govori koliko je dobar model.

A screenshot of a graph

Description automatically generated

Slika 15. ROC krivulja za varijablu "time".

## Analiza osjetljivosti

A screenshot of a computer

Description automatically generated

Postavljanje dokaza unutar Bayesovoj mreže (evidence).

### Evidence 1

Stavit ćemo da su 100% svi muškarci i promatrat ćemo što se događa s ciljnom varijablom.

A computer network with many different colored lines

Description automatically generated with medium confidence

Slika 16. Evidence 1

Iz ovog vidimo rezultate ciljne varijabla:

* 75% cenzura
* 25 % neuspjeh

### Evidence 2

Svi muškarci, bijelci, godine u rasponu od 34 do 70, heteroseksualci:

A computer network with many different colored lines

Description automatically generated with medium confidence

Slika 17. Evidence 2

Rezultati ciljne varijable:

* 84% cenzura
* 16 % neuspjeh

### Evidence 3

Pacijenti su sve žene, ne bijelkinje, između 69 i 74 kg, iznad 34 godine, cd40 u rasponu 0-235ml , cd420 u rasponu od 49-237 ml , i vrijeme trajanje terapije između 14-890 dana.

A computer network with many different colored boxes

Description automatically generated with medium confidence

Slika 18. Evidence 3

Rezultati ciljne varijable:

* 10% cenzura
* 90 % neuspjeh

### Evidence 4

Za razliku od prethodnog evidenc-a, sad ćemo povećati broj bijelih krvnih stanica cd40 i cd 420, a ostale varijable ćemo ostaviti iste.

A computer network with many different colored squares

Description automatically generated with medium confidence

Slika 19. Evidence 4

Rezultati ciljne varijable:

* 45% cenzura
* 55 % neuspjeh

Vidimo da veliki utjecaj imaju varijable cd40 i cd420 kako smo i izračunali.

### Evidence 5

Sad ćemo uključiti i varijablu **preanti** koja je bila četvrta jakosna IV.

A network of information on a white background

Description automatically generated with medium confidence

Slika 20. Evidence 5

Rezultati ciljne varijable:

* 45% cenzura
* 55 % neuspjeh

Možemo vidjeti da je ciljna varijabla ostala nepromijenjena.

### Evidence 6

Sad ćemo povećati vrijeme trajanja terapije između 890 I 1069 dana.

Rezultati ciljne varijable:

A computer network with many different colored lines

Description automatically generated with medium confidence

Slika 21. Evidence 6

* 88% cenzura
* 12 % neuspjeh

Vidljivo je da smo samo s prvim većim razredom varijable time postigli veliki rezultat, odnosno sa 55% neuspjeha pali smo na 12%, što pokazuje da ova varijabla ima najveći utjecan na ciljnu varijablu.