

Exercițiu laborator 11

Regresie logistică Bayesiană

O școală vrea să determine dacă un elev va promova sau nu un examen final, pe baza a două variabile predictive:

- număr ore de studiu pe săptămână;
- număr ore de somn pe zi.

În fișierul `date_promovare_examen.csv` aveți datele a 500 de studenți.

- a) Determinați dacă datele sunt balansate și construiți în PyMC un model de regresie logistică pentru această problemă.
- b) Care este în medie granița de decizie pentru acest model? Sunt datele bine separate?
- c) Ce credeți că influențează mai mult promovabilitatea, orele de studiu sau de somn?