

Problem 1. NaiveBayes

The data file, *data.csv*, contains observations on the impact of a number of treatment protocols on patients, where

Age is the age of the patient

Treatment is the treatment protocol (A, B or C)

Outcome describes the treatment as Ineffective(I), Good(G) or Excellent(E).

For this Q, let's assume/pretend that Age is normally distributed (so that no transformation is necessary.)

Design and code up a naive Bayes classifier to

[a] predict the probabilities that

treatment A on an 50-year old patient will turn out to be Good,

treatment A on an 50-year old patient will turn out to be Excellent

treatment B on an 50-year old patient will turn out to be Good,

treatment B on an 50-year old patient will turn out to be Excellent

treatment C on an 50-year old patient will turn out to be Good,

treatment C on an 50-year old patient will turn out to be Excellent

[b] identify the treatment you'd recommend for an x-year old patient. Briefly justify your recommendation.

You may find the following resources useful.

https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.GaussianNB.html

https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.MultinomialNB.html