Name: Punnal Ismail Khan **ID:21100019**

Report

Introduction:

In machine learning, naive Bayes classifiers are a family of simple "probabilistic classifiers" based on applying Bayes' theorem with strong independence assumptions between the features. I have used the Bernoulli naive Bayes model for this classification task because the Bernoulli model requires that all attributes value is binary (the dataset provided had binary values)

Training:

First I calculated the probability of classes i.e probability that the patient is normal or abnormal. Then I calculated individual probability with respect to each feature. For that I made a table for every test performed on the patient. The table contained 4 values: the probability of of getting positive value for a test given person is normal, the probability of of getting positive value for a test given person is abnormal, the probability of of getting negative value for a test given person is normal, and the probability of of getting negative value for a test given person is abnormal. Probability of getting a positive and negative value for every test is also computed. All the above computation was done on a training set.

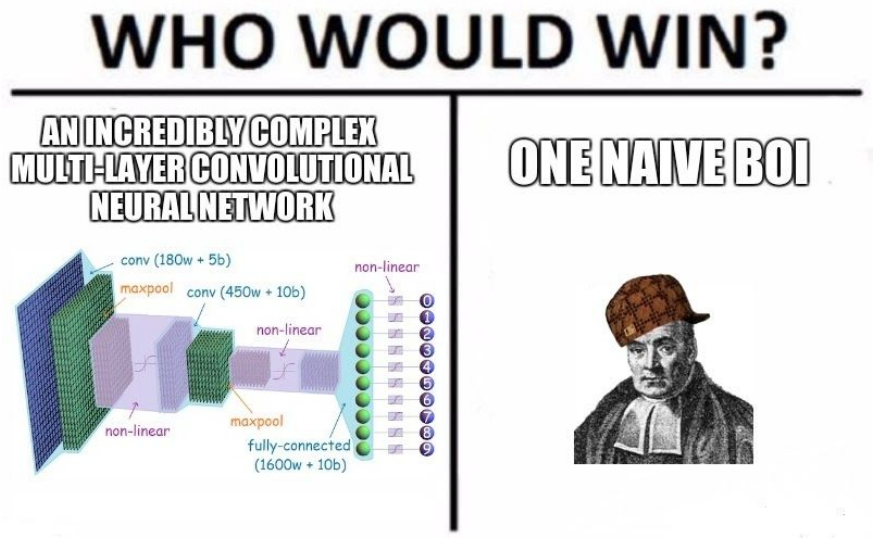
Testing:

Results of all the tests for the patients are taken(from testing data set) and the probability of getting a positive value on the test given person is normal is taken from tables computed during training for all tests. All probabilities are multiplied together to get a combined probability (let's call it a). Similarly, the probability of getting a positive value on the test given person is abnormal is taken from tables computed during training for all tests. All probabilities are multiplied together to get a combined probability (let's call it b). Then I divide both combined probabilities by evidence to normalize it. Evidence is the combined probability of getting positive(pre computing during training) on all tests. If combined probability a is greater than combined probability b, then I classify a person as normal. Otherwise I classify him as abnormal.

Accuracy:

Accuracy is calculated by comparing predicted value by original value. I get accuracy by the formula: (correctly predicted patients/total patients)*100

Meme:



Page complete yaaaay!!!