**Naive Bayes**

At first, we just applied a Gaussian Naive Bayes classifier without optimization, the accuracies got from training set and test set are both 0.66. Later, we quickly found that not all features are suitable for the Gaussian Naive Bayes classifiers. Because not all the features are numerical or continuous.

Specifically, all features in this dataset can be divided into three types: numerical (or continuous), binary, and categorical. In the level of fields:

*# age in days /numerical*

*# BMI (from height and weight) /numerical*

*# MAP (from ap\_hi and ap\_lo) /numerical*

*# cholesterol 1: normal, 2: above normal, 3: well above normal /categorical*

*# gluc 1: normal, 2: above normal, 3: well above normal /categorical*

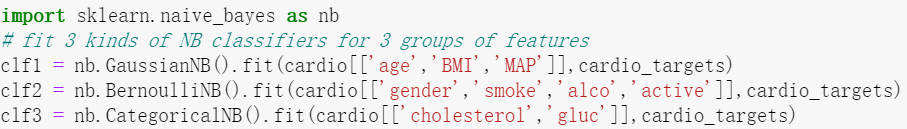
*# gender 1 - women, 2 - men /binary*

*# smoke whether patient smokes or not /binary*

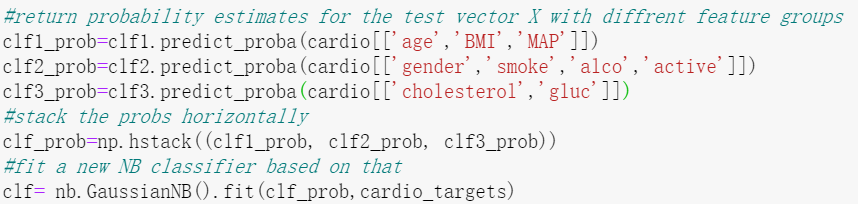
*# alco Binary feature /binary*

*# active Binary feature /binary*

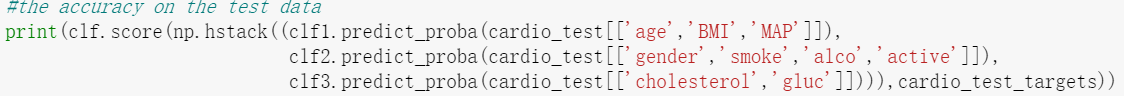
Given these 3 types of features, we can make different assumptions regarding the distribution of the likelihood, i.e., Gaussian, Bernoulli, and Categorical distribution. And we fit 3 Naïve Bayesian classifiers to the training datasets sliced by the corresponding fields as sub-classifiers:



In order to assemble these three sub-classifiers into the final classifier, we calculated the probabilities estimate for the sliced training vectors. Then the probabilities were stacked horizontally for the fitting of a new Gaussian Naïve Bayesian classifier, since the stacked probabilities are all numerical fields. And we got the final Naïve Bayesian classifier after fitting it on the stacked probabilities and the original target field.



We calculated the accuracies for the training and test set. They are both improved to 0.68. It is worth mentioning that the accuracy of the test set is also calculated by stacking the probabilities returned by the three sub-classifiers:



Naive Bayes learners and classifiers can be relative fast compared to more sophisticated methods. On the other hand, naïve Bayes is known to be a bad estimator, so the probability outputs from predict\_proba are not to be taken too seriously. Additionally, there is no ‘zero probability problem’ in this binary classed classification problem because both classes appeared in the training data.