Marc Wafik William 4141127

***Show that Bayes classifier is the best classifier that could be used***

Usually, a dataset is considered to consist of samples of a distribution that generates your data. Then, you build a predictive model from the given data: given a sample, you predict the class f^(xi)f^(xi), whereas the real class of the sample is f(xi)f(xi).

However, in theory, you could decide not to choose one particular model but rather consider all possible models at once and combine them somehow into one big model F^F^.

Of course, given the data, many of the smaller models could be quite improbable or inappropriate (for example, models that predict only one value of the target, even though there are multiple values of the target in your dataset).

In any case, you want to predict the target value of new samples, which are drawn from the same distribution

Using Bayes formula, you can compute, what is the probability that a new sample xx has target value vv, given the data DD:

P(v∣D)=∑f^P(v∣f^)P(f^∣D).

***What is the problem with this classifier?***

In classification tasks, you need a big data set in order to make reliable estimations of the probability of each class. You can use Naïve Bayes classification algorithm with a small data set but precision and recall will keep very low.