3.1.2 Naïve Bayes

We choose a BernoulliNB classifier, because data in most of our features is distributed according to multivariate Bernoulli distributions; i.e., there may be multiple features but each one is assumed to be a binary-valued feature vectors. This is the result of one-of-K encoding during data preprocessing.