

This assignment using Gaussian Naïve Bayes classifier to finish. In this task, the biggest problem encountered is that the variance of some variables is 0. In the final result, we use 0.00001 as the variance of a variable which variance is 0. We also tried to use discrete prior probabilities to treat variables with a variance of 0 as Bernoulli distributions. Treat variable values greater than the mean as 1 and less than the mean as 0. And make the probability non-zero through Laplace transform, though we are not reaching the better result.

In the final result, our naïve Bayes algorithm achieves 80% accuracy, and the Precision is 60.7%, Recall is 1, F1-score is 76%. Compare Naïve Bayes of sklearn, my algorithm is much worse. I think the flaw is mainly in the way of dealing with attribute which variance is zero.