	Naive Bayes	SGD
train acc	83.7%	92.6%
test acc	84.5%	91.8 %

Table 1: Training and testing accuracy for Naive Bayes and SGD

For the Naive Bayes method, we found that the training accuracy is higher than the test accuracy. This could be the fact that there are more noise in the training data comparing to the testing data and Naive Bayes model are more affected by the noise. For stochastic descent, we noticed that the training accuracy is higher than the testing accuracy. This could due to the model tends to overfit the training data and then performs not as well on the testing data.

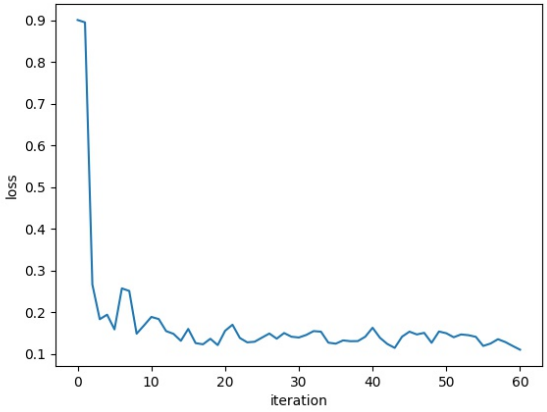

Figure 1: Training loss with learning rate 0.2

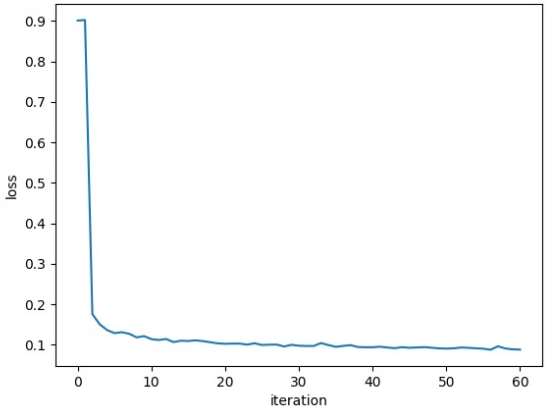

Figure 2: Training loss with learning rate 0.005

We can see both learning rate at 0.2 and learning rate at 0.005 converges to low loss. We noticed that higher learning rate will have more loss fluctuation comparing to the low learning rate. This is because when the learning rate is high, it is more likely that the answer is jumping back and forth between the minimization point. With low learning rate, the answer move to the minimal point slowly, resulting in constantly decrease in loss.