DAT565/DIT407 Assignment 6

Sebastian Miles
miless@chalmers.se

Olle Lapidus
ollelap@chalmers.se

2024-10-11

Problem 1

We verify that the images are 28x28 pixels grayscales and plot a few of the images.

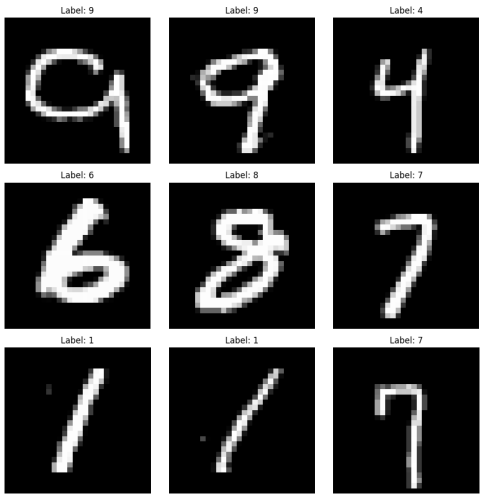

Figure 1: Various images from the datasets

Problem 2

We train the data with a single hidden layer. We put the hidden layer size as 128 and the learning rate for the SGD as 0.01. The accuracy over 10 epochs is shown in figure 1.

Problem 3

MultinomialNB

See table 2
Precision (ham): 0.95

Epoch	Accuracy
1	79.45%
2	82.06%
3	84.35%
4	83.96%
5	84.61%
6	85.36%
7	84.10%
8	85.57%
9	84.92%
10	86.75%

Table 1: Accuracy of the test data over 10 epochs.

	Predicted Ham	Predicted Spam	Sum
Actual Ham	1281	1	1282
Actual Spam	73	171	244
Sum	1354	172	1526

Table 2: Confusion Matrix for MultinomialNB

Precision (spam): 0.99
Recall (ham): 1.00
Recall (spam): 0.70

BernoulliNB

	Predicted Ham	Predicted Spam	Sum
Actual Ham	1280	2	1282
Actual Spam	180	64	244
Sum	1460	68	1528

Table 3: Confusion Matrix for BernoulliNB

See table 3
Precision (ham): 0.88
Precision (spam): 0.97
Recall (ham): 1.00
Recall (spam): 0.26

The classifier demonstrates high accuracy, with a precision of 0.99 for spam detection, although the recall for spam is somewhat lower, indicating occasional missed spam emails.

Problem 4

Code