Training dataset size: 50000
Test dataset size: 10000
Start by dividing the data into subsets and applying the transform to these subsets
Training subset size: 5000
Test subset size: 1000
Features loaded from saved files.
The Gaussian Naive-Bayes algorithm that was manually implemented:
From-scratch model loaded.
Model has been fit
Predictions have been set
Now evaluating the Gaussian Naive-Bayes algorithm that was manually implemented:
Accuracy: 77.30%
Confusion Matrix (rows = true labels 0-9, columns = predictions):
[[79 1 1 1 0 0 1 0 14 3]
[4 86 0 2 0 0 0 1 1 6]
[8 0 59 9 11 2 10 0 1 0]
[1 0 3 75 2 9 8 1 1 0]
[1 0 6 6 75 4 3 5 0 0]
[0 1 6 16 2 70 3 2 0 0]
[2 0 6 5 6 1 79 1 0 0]
[2 1 0 5 9 6 0 76 0 1]
[8 0 2 0 0 0 0 0 87 3]
[5 3 0 1 0 0 0 1 3 87]]
Classification Report:

	precision	recall	f1-score	support
0	0.72	0.79	0.75	100
1	0.93	0.86	0.90	100
2	0.71	0.59	0.64	100
3	0.62	0.75	0.68	100
4	0.71	0.75	0.73	100
5	0.76	0.70	0.73	100
6	0.76	0.79	0.77	100
7	0.87	0.76	0.81	100
8	0.81	0.87	0.84	100
9	0.87	0.87	0.87	100
accuracy			0.77	1000
macro avg	0.78	0.77	0.77	1000
weighted avg	0.78	0.77	0.77	1000

The SciKit Gaussian Naive-Bayes algorithm:
Scikit model loaded.
Model has been fit

Predictions have been set
Now evaluating the SciKit Gaussian Naive-Bayes algorithm:
Scikit-learn GaussianNB Accuracy: 77.30%
Confusion Matrix (rows = true labels 0-9, columns = predictions):
[[79 1 1 1 0 0 1 0 14 3]
[4 86 0 2 0 0 0 1 1 6]
[8 0 59 9 11 2 10 0 1 0]
[1 0 3 75 2 9 8 1 1 0]
[1 0 6 6 75 4 3 5 0 0]
[0 1 6 16 2 70 3 2 0 0]
[2 0 6 5 6 1 79 1 0 0]
[2 1 0 5 9 6 0 76 0 1]
[8 0 2 0 0 0 0 0 87 3]
[5 3 0 1 0 0 0 1 3 87]]
Classification Report:

	precision	recall	f1-score	support
0	0.72	0.79	0.75	100
1	0.93	0.86	0.90	100
2	0.71	0.59	0.64	100
3	0.62	0.75	0.68	100
4	0.71	0.75	0.73	100
5	0.76	0.70	0.73	100
6	0.76	0.79	0.77	100
7	0.87	0.76	0.81	100
8	0.81	0.87	0.84	100
9	0.87	0.87	0.87	100
accuracy			0.77	1000
macro avg	0.78	0.77	0.77	1000
weighted avg	0.78	0.77	0.77	1000