MAX DEPTH = 50

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 Decision Tree algorithm that was manually implemented:
Loaded saved decision tree model
Model has been fit
Predictions have been set
Now evaluating the Decision Tree algorithm that was manually implemented:
Accuracy: 58.60%
Confusion Matrix (rows = true labels 0-9, columns = predictions):
[[59 5 9 1 2 0 2 2 14 6]
[4 70 1 2 1 0 2 1 7 12]
[6 1 44 10 12 8 14 3 2 0]
[2 2 7 41 4 22 10 9 2 1]
[4 0 10 5 49 9 4 18 0 1]
[0 0 4 22 6 54 5 5 4 0]
[4 0 10 4 6 1 71 3 0 1]
[2 1 5 6 9 11 1 61 3 1]
[21 4 2 1 0 0 1 1 62 8]
[5 10 0 1 1 0 0 0 8 75]]
Classification Report:

	precision	recall	f1-score	support
0	0.55	0.59	0.57	100
1	0.75	0.70	0.73	100
2	0.48	0.44	0.46	100
3	0.44	0.41	0.42	100
4	0.54	0.49	0.52	100
5	0.51	0.54	0.53	100
6	0.65	0.71	0.68	100
7	0.59	0.61	0.60	100
8	0.61	0.62	0.61	100
9	0.71	0.75	0.73	100
accuracy			0.59	1000
macro avg	0.58	0.59	0.58	1000
weighted avg	0.58	0.59	0.58	1000

The SciKit Decision Tree algorithm:

The Decision Tree algorithm using Scikit-learn:
Loaded saved decision tree model
Model has been fit
Predictions have been set
Now evaluating the Decision Tree algorithm using Scikit-learn:
Accuracy: 58.30%
Confusion Matrix (rows = true labels 0-9, columns = predictions):
[[57 4 7 2 3 0 2 3 17 5]
[4 66 1 3 1 0 1 0 9 15]
[3 0 47 12 8 8 16 5 1 0]
[1 1 11 44 5 17 11 7 1 2]
[2 1 9 5 50 12 4 16 1 0]
[1 0 6 18 7 55 5 6 2 0]
[3 0 15 5 5 2 69 1 0 0]
[1 0 4 9 10 13 1 59 1 2]
[21 5 2 1 0 0 1 1 63 6]
[9 9 0 1 1 0 0 0 7 73]]
Classification Report:

	precision	recall	f1-score	support
0	0.56	0.57	0.56	100
1	0.77	0.66	0.71	100
2	0.46	0.47	0.47	100
3	0.44	0.44	0.44	100
4	0.56	0.50	0.53	100
5	0.51	0.55	0.53	100
6	0.63	0.69	0.66	100
7	0.60	0.59	0.60	100
8	0.62	0.63	0.62	100
9	0.71	0.73	0.72	100
accuracy			0.58	1000
macro avg	0.59	0.58	0.58	1000
weighted avg	0.59	0.58	0.58	1000

MAX DEPTH = 10

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 Decision Tree algorithm that was manually implemented:

Loaded saved decision tree model
Model has been fit
Predictions have been set
Now evaluating the Decision Tree algorithm that was manually implemented:
Accuracy: 61.40%
Confusion Matrix (rows = true labels 0-9, columns = predictions):
[[59 6 8 3 3 1 0 1 15 4]
[3 71 1 2 1 1 2 1 6 12]
[4 1 43 18 8 13 11 1 1 0]
[1 0 8 54 3 21 9 3 0 1]
[4 0 7 6 59 6 2 15 0 1]
[0 0 4 24 4 60 2 5 1 0]
[5 0 9 11 2 2 69 2 0 0]
[1 1 5 12 11 8 1 60 1 0]
[19 4 2 0 0 0 2 2 64 7]
[5 9 0 1 1 0 1 0 8 75]]
Classification Report:

	precision	recall	f1-score	support
0	0.58	0.59	0.59	100
1	0.77	0.71	0.74	100
2	0.49	0.43	0.46	100
3	0.41	0.54	0.47	100
4	0.64	0.59	0.61	100
5	0.54	0.60	0.57	100
6	0.70	0.69	0.69	100
7	0.67	0.60	0.63	100
8	0.67	0.64	0.65	100
9	0.75	0.75	0.75	100
accuracy			0.61	1000
macro avg	0.62	0.61	0.62	1000
weighted avg	0.62	0.61	0.62	1000

The SciKit Decision Tree algorithm:
The Decision Tree algorithm using Scikit-learn:
Loaded saved decision tree model
Model has been fit
Predictions have been set
Now evaluating the Decision Tree algorithm using Scikit-learn:
Accuracy: 60.80%
Confusion Matrix (rows = true labels 0-9, columns = predictions):
[[60 4 6 3 2 1 0 1 18 5]
[3 68 1 4 1 1 1 1 7 13]

[3 0 42 17 7 13 13 3 2 0]
[2 0 8 57 3 18 8 3 0 1]
[4 0 6 7 58 9 2 12 1 1]
[0 0 6 21 4 61 3 4 1 0]
[3 0 13 12 3 3 65 1 0 0]
[1 0 3 12 13 13 0 58 0 0]
[18 3 2 0 1 0 2 3 66 5]
[6 8 0 3 2 0 1 0 7 73]]
Classification Report:

	precision	recall	f1-score	support
0	0.60	0.60	0.60	100
1	0.82	0.68	0.74	100
2	0.48	0.42	0.45	100
3	0.42	0.57	0.48	100
4	0.62	0.58	0.60	100
5	0.51	0.61	0.56	100
6	0.68	0.65	0.67	100
7	0.67	0.58	0.62	100
8	0.65	0.66	0.65	100
9	0.74	0.73	0.74	100
accuracy			0.61	1000
macro avg	0.62	0.61	0.61	1000
weighted avg	0.62	0.61	0.61	1000

MAX DEPTH = 25

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 Decision Tree algorithm that was manually implemented:
Loaded saved decision tree model
Model has been fit
Predictions have been set
Now evaluating the Decision Tree algorithm that was manually implemented:
Accuracy: 58.60%
Confusion Matrix (rows = true labels 0-9, columns = predictions):
[[59 5 9 1 2 0 2 2 14 6]
[4 70 1 2 1 0 2 1 7 12]
[6 1 44 10 12 8 14 3 2 0]

[2 2 7 41 4 22 10 9 2 1]
[4 0 10 5 49 9 4 18 0 1]
[0 0 4 22 6 54 5 5 4 0]
[4 0 10 4 6 1 71 3 0 1]
[2 1 5 6 9 11 1 61 3 1]
[21 4 2 1 0 0 1 1 62 8]
[5 10 0 1 1 0 0 0 8 75]]
Classification Report:

	precision	recall	f1-score	support
0	0.55	0.59	0.57	100
1	0.75	0.70	0.73	100
2	0.48	0.44	0.46	100
3	0.44	0.41	0.42	100
4	0.54	0.49	0.52	100
5	0.51	0.54	0.53	100
6	0.65	0.71	0.68	100
7	0.59	0.61	0.60	100
8	0.61	0.62	0.61	100
9	0.71	0.75	0.73	100
accuracy			0.59	1000
macro avg	0.58	0.59	0.58	1000
weighted avg	0.58	0.59	0.58	1000

The SciKit Decision Tree algorithm:
The Decision Tree algorithm using Scikit-learn:
Loaded saved decision tree model
Model has been fit
Predictions have been set
Now evaluating the Decision Tree algorithm using Scikit-learn:
Accuracy: 58.30%
Confusion Matrix (rows = true labels 0-9, columns = predictions):
[[57 4 7 2 3 0 2 3 17 5]
[4 66 1 3 1 0 1 0 9 15]
[3 0 47 12 8 8 16 5 1 0]
[1 1 11 44 5 17 11 7 1 2]
[2 1 9 5 50 12 4 16 1 0]
[1 0 6 18 7 55 5 6 2 0]
[3 0 15 5 5 2 69 1 0 0]
[1 0 4 9 10 13 1 59 1 2]
[21 5 2 1 0 0 1 1 63 6]
[9 9 0 1 1 0 0 0 7 73]]
Classification Report:

	precision	recall	f1-score	support
0	0.56	0.57	0.56	100
1	0.77	0.66	0.71	100
2	0.46	0.47	0.47	100
3	0.44	0.44	0.44	100
4	0.56	0.50	0.53	100
5	0.51	0.55	0.53	100
6	0.63	0.69	0.66	100
7	0.60	0.59	0.60	100
8	0.62	0.63	0.62	100
9	0.71	0.73	0.72	100
accuracy			0.58	1000
macro avg	0.59	0.58	0.58	1000
weighted avg	0.59	0.58	0.58	1000

MAX DEPTH = 75

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 Decision Tree algorithm that was manually implemented:
Loaded saved decision tree model
Model has been fit
Predictions have been set
Now evaluating the Decision Tree algorithm that was manually implemented:
Accuracy: 58.60%
Confusion Matrix (rows = true labels 0-9, columns = predictions):
[[59 5 9 1 2 0 2 2 14 6]
[4 70 1 2 1 0 2 1 7 12]
[6 1 44 10 12 8 14 3 2 0]
[2 2 7 41 4 22 10 9 2 1]
[4 0 10 5 49 9 4 18 0 1]
[0 0 4 22 6 54 5 5 4 0]
[4 0 10 4 6 1 71 3 0 1]
[2 1 5 6 9 11 1 61 3 1]
[21 4 2 1 0 0 1 1 62 8]
[5 10 0 1 1 0 0 0 8 75]]
Classification Report:

	precision	recall	f1-score	support

0	0.55	0.59	0.57	100
1	0.75	0.70	0.73	100
2	0.48	0.44	0.46	100
3	0.44	0.41	0.42	100
4	0.54	0.49	0.52	100
5	0.51	0.54	0.53	100
6	0.65	0.71	0.68	100
7	0.59	0.61	0.60	100
8	0.61	0.62	0.61	100
9	0.71	0.75	0.73	100
accuracy			0.59	1000
macro avg	0.58	0.59	0.58	1000
weighted avg	0.58	0.59	0.58	1000

The SciKit Decision Tree algorithm:
The Decision Tree algorithm using Scikit-learn:
Loaded saved decision tree model
Model has been fit
Predictions have been set
Now evaluating the Decision Tree algorithm using Scikit-learn:
Accuracy: 58.30%
Confusion Matrix (rows = true labels 0-9, columns = predictions):
[[57 4 7 2 3 0 2 3 17 5]
[4 66 1 3 1 0 1 0 9 15]
[3 0 47 12 8 8 16 5 1 0]
[1 1 11 44 5 17 11 7 1 2]
[2 1 9 5 50 12 4 16 1 0]
[1 0 6 18 7 55 5 6 2 0]
[3 0 15 5 5 2 69 1 0 0]
[1 0 4 9 10 13 1 59 1 2]
[21 5 2 1 0 0 1 1 63 6]
[9 9 0 1 1 0 0 0 7 73]]
Classification Report:

	precision	recall	f1-score	support
0	0.56	0.57	0.56	100
1	0.77	0.66	0.71	100
2	0.46	0.47	0.47	100
3	0.44	0.44	0.44	100
4	0.56	0.50	0.53	100
5	0.51	0.55	0.53	100
6	0.63	0.69	0.66	100

7	0.60	0.59	0.60	100
8	0.62	0.63	0.62	100
9	0.71	0.73	0.72	100
accuracy			0.58	1000
macro avg	0.59	0.58	0.58	1000
weighted avg	0.59	0.58	0.58	1000

MAX DEPTH = 90

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 Decision Tree algorithm that was manually implemented:
Loaded saved decision tree model
Model has been fit
Predictions have been set
Now evaluating the Decision Tree algorithm that was manually implemented:
Accuracy: 58.60%
Confusion Matrix (rows = true labels 0-9, columns = predictions):
[[59 5 9 1 2 0 2 2 14 6]
[4 70 1 2 1 0 2 1 7 12]
[6 1 44 10 12 8 14 3 2 0]
[2 2 7 41 4 22 10 9 2 1]
[4 0 10 5 49 9 4 18 0 1]
[0 0 4 22 6 54 5 5 4 0]
[4 0 10 4 6 1 71 3 0 1]
[2 1 5 6 9 11 1 61 3 1]
[21 4 2 1 0 0 1 1 62 8]
[5 10 0 1 1 0 0 0 8 75]]
Classification Report:

	precision	recall	f1-score	support
0	0.55	0.59	0.57	100
1	0.75	0.70	0.73	100
2	0.48	0.44	0.46	100
3	0.44	0.41	0.42	100
4	0.54	0.49	0.52	100
5	0.51	0.54	0.53	100
6	0.65	0.71	0.68	100
7	0.59	0.61	0.60	100

8	0.61	0.62	0.61	100
9	0.71	0.75	0.73	100
accuracy			0.59	1000
macro avg	0.58	0.59	0.58	1000
weighted avg	0.58	0.59	0.58	1000

The SciKit Decision Tree algorithm:
The Decision Tree algorithm using Scikit-learn:
Loaded saved decision tree model
Model has been fit
Predictions have been set
Now evaluating the Decision Tree algorithm using Scikit-learn:
Accuracy: 58.30%
Confusion Matrix (rows = true labels 0-9, columns = predictions):

```
[[57  4  7  2  3  0  2  3 17  5]
 [ 4 66  1  3  1  0  1  0  9 15]
 [ 3  0 47 12  8  8 16  5  1  0]
 [ 1  1 11 44  5 17 11  7  1  2]
 [ 2  1  9  5 50 12  4 16  1  0]
 [ 1  0  6 18  7 55  5  6  2  0]
 [ 3  0 15  5  5  2 69  1  0  0]
 [ 1  0  4  9 10 13  1 59  1  2]
 [21  5  2  1  0  0  1  1 63  6]
 [ 9  9  0  1  1  0  0  0  7 73]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.56	0.57	0.56	100
1	0.77	0.66	0.71	100
2	0.46	0.47	0.47	100
3	0.44	0.44	0.44	100
4	0.56	0.50	0.53	100
5	0.51	0.55	0.53	100
6	0.63	0.69	0.66	100
7	0.60	0.59	0.60	100
8	0.62	0.63	0.62	100
9	0.71	0.73	0.72	100
accuracy			0.58	1000
macro avg	0.59	0.58	0.58	1000
weighted avg	0.59	0.58	0.58	1000