Name: Abdirahim Ahmed
Student ID: 801338570

Homework 6

Github repository:https://github.com/Abdiirahim/ECGR-4105-Intro-to-ML/tree/main/homewprk6

Problem 1 (30 pts):
a. Build a fully connected neural network for the housing dataset you did in previous homework. For training and validation use 80% (training) and 20% (validation) split. For this part, only use one hidden layer with 8 nodes. Train your network for as many epochs you want to. Report your training time, training loss, and evaluation accuracy for many epochs Analyze your results in your report. Make sure to submit your code by providing the GitHub URL of your course repository for this course. Compare your results against the linear regression and support vector regression from previous homework. (15pts)

A fully connected neural network with a single hidden layer containing 8 nodes was implemented to predict house prices. The features were normalized, and the model was trained using the Mean Squared Error (MSE) loss and the SGD optimizer. Training and validation losses were monitored over 1000 epochs, showing steady improvement and good generalization.

Epoch 100/1000 - Train Loss: 0.0240, Validation Loss: 0.0229
Epoch 200/1000 - Train Loss: 0.0175, Validation Loss: 0.0174
Epoch 300/1000 - Train Loss: 0.0152, Validation Loss: 0.0154
Epoch 400/1000 - Train Loss: 0.0139, Validation Loss: 0.0143
Epoch 500/1000 - Train Loss: 0.0129, Validation Loss: 0.0137
Epoch 600/1000 - Train Loss: 0.0121, Validation Loss: 0.0131
Epoch 700/1000 - Train Loss: 0.0113, Validation Loss: 0.0126
Epoch 800/1000 - Train Loss: 0.0105, Validation Loss: 0.0120
Epoch 900/1000 - Train Loss: 0.0096, Validation Loss: 0.0113
Epoch 1000/1000 - Train Loss: 0.0085, Validation Loss: 0.0104
Final Train Loss: 0.008531803265213966
Final Validation Loss: 0.01039207261055708

Problem 1a Analysis:
The single-hidden-layer model achieved a final train loss of 0.0085 and a final validation loss of 0.0104 after 1000 epochs. Both losses decreased steadily throughout training, with the validation loss closely following the train loss, indicating good generalization. The model effectively captured the relationships in the housing dataset without overfitting, demonstrating its capability to perform well on unseen data.

b. Extend your network with two more additional hidden layers, like the example we did in lecture. Train your network for many epochs as needed. Report your training time, training loss,

Name: Abdirahim Ahmed
Student ID: 801338570

and evaluation accuracy. Analyze your results in your report. Make sure to submit your code by providing the GitHub URL of your course repository for this course. Analyze your results in your report and compare your model size and accuracy over the baseline implementation in Problem1. a. Do you see any over-fitting? Compare your results against the linear regression and support vector regression from previous homework. Make sure to submit your code by providing the GitHub URL of your course repository for this course. (25pts)

The network was extended to include three hidden layers with 64, 32, and 16 nodes, respectively, to increase the model's capacity. Dropout regularization was introduced to prevent overfitting, and the Adam optimizer was used for faster convergence. The additional layers improved the validation loss, but the training time was slightly longer due to the model's complexity

Epoch 100/1000 - Train Loss: 0.0975, Validation Loss: 0.0274
Epoch 200/1000 - Train Loss: 0.0584, Validation Loss: 0.0096
Epoch 300/1000 - Train Loss: 0.0475, Validation Loss: 0.0063
Epoch 400/1000 - Train Loss: 0.0480, Validation Loss: 0.0055
Epoch 500/1000 - Train Loss: 0.0522, Validation Loss: 0.0064
Epoch 600/1000 - Train Loss: 0.0445, Validation Loss: 0.0049
Epoch 700/1000 - Train Loss: 0.0482, Validation Loss: 0.0044
Epoch 800/1000 - Train Loss: 0.0465, Validation Loss: 0.0061
Epoch 900/1000 - Train Loss: 0.0506, Validation Loss: 0.0045
Epoch 1000/1000 - Train Loss: 0.0376, Validation Loss: 0.0063
Final Train Loss: 0.037569064646959305

Problem 1b Analysis:
The three-hidden-layer model achieved a final train loss of 0.0376 and a final validation loss of 0.0063 after 1000 epochs. The deeper network demonstrated improved validation performance compared to the single-layer model, likely due to its increased capacity to model complex relationships. The use of dropout regularization helped prevent overfitting, as evidenced by the validation loss remaining lower than the train loss. This model is better suited for applications where validation performance is critical.

Problem 2 (40 pts):
a. Use the cancer dataset to build a fully connected neural network to classify the type of cancer (Malignant vs. benign). For training and validation use 80% (training) and 20% (validation) split. For this part, only use one hidden layer with 32 nodes. Train your network for many epochs as needed. Report your training time, training loss, and evaluation accuracy. Analyze your results in your report. Compare your results against the logistic regression and support vector classification from previous homework. Make sure to submit your code by providing the GitHub URL of your course repository for this course. (15pts)

Name: Abdirahim Ahmed
Student ID: 801338570

A fully connected neural network with a single hidden layer of 32 nodes was used to classify cancer as malignant or benign. The model was trained with the Binary Cross-Entropy (BCE) loss and the Adam optimizer. The features were normalized, and the model showed a steady decrease in loss over 1000 epochs. Slight overfitting was observed in the later epochs, but the overall generalization was acceptable.

Epoch 100/1000 - Train Loss: 0.1647, Validation Loss: 0.1503
Epoch 200/1000 - Train Loss: 0.0857, Validation Loss: 0.0794
Epoch 300/1000 - Train Loss: 0.0616, Validation Loss: 0.0651
Epoch 400/1000 - Train Loss: 0.0495, Validation Loss: 0.0615
Epoch 500/1000 - Train Loss: 0.0412, Validation Loss: 0.0603
Epoch 600/1000 - Train Loss: 0.0347, Validation Loss: 0.0606
Epoch 700/1000 - Train Loss: 0.0290, Validation Loss: 0.0621
Epoch 800/1000 - Train Loss: 0.0239, Validation Loss: 0.0642
Epoch 900/1000 - Train Loss: 0.0195, Validation Loss: 0.0667
Epoch 1000/1000 - Train Loss: 0.0159, Validation Loss: 0.0697
Final Train Loss: 0.01588340476155281
Final Validation Loss: 0.06972919404506683

Problem 2a Analysis:
The single-hidden-layer model achieved a final train loss of 0.0159 and a final validation loss of 0.0697 after 1000 epochs. Both losses steadily decreased during training, with the validation loss showing slight degradation toward the end, suggesting possible slight overfitting. The model performed well on the cancer dataset, effectively separating the malignant and benign classes. The validation loss remained reasonably low, indicating good generalization for a single-hidden-layer network.

b. Extend your network with two more additional hidden layers, like the example we did in lecture. Train your network for many epochs as needed. Report your training time, training loss, and evaluation accuracy for many epochs as needed. Analyze your results in your report. Make sure to submit your code by providing the GitHub URL of your course repository for this course. Analyze your results in your report and compare your model size and accuracy over the baseline implementation in Problem1. a. Do you see any over-fitting? Compare your results against the logistic regression and support vector classification from previous homework. Make sure to submit your code by providing the GitHub URL of your course repository for this course. (25pts)

The network was extended to three hidden layers with 64, 32, and 16 nodes. Despite achieving an extremely low training loss, the validation loss began increasing after 200 epochs, indicating significant overfitting. The BCE loss and Adam optimizer were used, and dropout regularization is recommended to improve generalization further

Name: Abdirahim Ahmed
Student ID: 801338570

Epoch 100/1000 - Train Loss: 0.0477, Validation Loss: 0.0601
Epoch 200/1000 - Train Loss: 0.0136, Validation Loss: 0.0695
Epoch 300/1000 - Train Loss: 0.0016, Validation Loss: 0.0906
Epoch 400/1000 - Train Loss: 0.0005, Validation Loss: 0.1075
Epoch 500/1000 - Train Loss: 0.0003, Validation Loss: 0.1200
Epoch 600/1000 - Train Loss: 0.0002, Validation Loss: 0.1297
Epoch 700/1000 - Train Loss: 0.0001, Validation Loss: 0.1372
Epoch 800/1000 - Train Loss: 0.0001, Validation Loss: 0.1436
Epoch 900/1000 - Train Loss: 0.0001, Validation Loss: 0.1493
Epoch 1000/1000 - Train Loss: 0.0000, Validation Loss: 0.1543
Final Train Loss: 4.543393515632488e-05
Final Validation Loss: 0.1542828530073166

Problem 2b Analysis:
The three-hidden-layer model achieved a final train loss of 4.54e-05 and a final validation loss of 0.1543 after 1000 epochs. While the train loss was reduced to a very low value, the validation loss began increasing after the first 200 epochs. This is a classic sign of overfitting, where the model memorized the training data but struggled to generalize to unseen data. The additional layers increased the model's capacity but also made it more prone to overfitting, requiring techniques like early stopping or dropout for better performance.

Problem 3 (40 pts):
a. Create a fully connected Neural Network for all 10 classes in CIFAR-10 with only one hidden layer with the size of 256. Train your network for 100 epochs. Report your training time, training loss and evaluation accuracy after 100 epochs. Analyze your results in your report. Make sure to submit your code by providing the GitHub URL of your course repository for this course. (25pt)
fully connected neural network with a single hidden layer of 256 nodes was implemented to classify CIFAR-10 images into 10 classes. The model was trained using the Cross-Entropy loss and the SGD optimizer. The network achieved moderate accuracy (~50%) but struggled to capture spatial relationships in the image data due to its fully connected architecture

Epoch 1/100, Loss: 1.8846, Accuracy: 40.32%
Epoch 2/100, Loss: 1.6635, Accuracy: 43.82%
Epoch 3/100, Loss: 1.5794, Accuracy: 45.95%
Epoch 4/100, Loss: 1.5239, Accuracy: 46.78%
Epoch 5/100, Loss: 1.4802, Accuracy: 48.32%
Epoch 6/100, Loss: 1.4406, Accuracy: 48.96%
Epoch 7/100, Loss: 1.4071, Accuracy: 49.48%
Epoch 8/100, Loss: 1.3752, Accuracy: 49.92%
Epoch 9/100, Loss: 1.3459, Accuracy: 50.40%

Name: Abdirahim Ahmed
Student ID: 801338570

Epoch 10/100, Loss: 1.3188, Accuracy: 51.25%
Epoch 11/100, Loss: 1.2936, Accuracy: 51.00%
Epoch 12/100, Loss: 1.2684, Accuracy: 51.75%
Epoch 13/100, Loss: 1.2459, Accuracy: 52.10%
Epoch 14/100, Loss: 1.2237, Accuracy: 52.32%
Epoch 15/100, Loss: 1.2033, Accuracy: 53.08%
Epoch 16/100, Loss: 1.1840, Accuracy: 51.88%
Epoch 17/100, Loss: 1.1636, Accuracy: 52.56%
Epoch 18/100, Loss: 1.1452, Accuracy: 52.65%
Epoch 19/100, Loss: 1.1279, Accuracy: 52.96%
Epoch 20/100, Loss: 1.1095, Accuracy: 52.47%
Epoch 21/100, Loss: 1.0912, Accuracy: 51.80%
Epoch 22/100, Loss: 1.0742, Accuracy: 53.18%
Epoch 23/100, Loss: 1.0574, Accuracy: 53.05%
Epoch 24/100, Loss: 1.0416, Accuracy: 52.91%
Epoch 25/100, Loss: 1.0251, Accuracy: 53.27%
Epoch 26/100, Loss: 1.0094, Accuracy: 53.42%
Epoch 27/100, Loss: 0.9932, Accuracy: 53.71%
Epoch 28/100, Loss: 0.9779, Accuracy: 53.48%
Epoch 29/100, Loss: 0.9623, Accuracy: 52.65%
Epoch 30/100, Loss: 0.9479, Accuracy: 53.19%
Epoch 31/100, Loss: 0.9322, Accuracy: 53.31%
Epoch 32/100, Loss: 0.9174, Accuracy: 52.66%
Epoch 33/100, Loss: 0.9024, Accuracy: 53.59%
Epoch 34/100, Loss: 0.8899, Accuracy: 53.13%
Epoch 35/100, Loss: 0.8748, Accuracy: 53.58%
Epoch 36/100, Loss: 0.8607, Accuracy: 53.43%
Epoch 37/100, Loss: 0.8467, Accuracy: 53.04%
Epoch 38/100, Loss: 0.8326, Accuracy: 52.97%
Epoch 39/100, Loss: 0.8200, Accuracy: 53.24%
Epoch 40/100, Loss: 0.8075, Accuracy: 52.86%
Epoch 41/100, Loss: 0.7925, Accuracy: 53.07%
Epoch 42/100, Loss: 0.7812, Accuracy: 53.19%
Epoch 43/100, Loss: 0.7668, Accuracy: 52.93%
Epoch 44/100, Loss: 0.7544, Accuracy: 52.12%
Epoch 45/100, Loss: 0.7431, Accuracy: 52.61%
Epoch 46/100, Loss: 0.7314, Accuracy: 52.89%
Epoch 47/100, Loss: 0.7158, Accuracy: 51.69%
Epoch 48/100, Loss: 0.7055, Accuracy: 51.56%
Epoch 49/100, Loss: 0.6931, Accuracy: 52.95%
Epoch 50/100, Loss: 0.6832, Accuracy: 51.60%
Epoch 51/100, Loss: 0.6679, Accuracy: 51.10%
Epoch 52/100, Loss: 0.6586, Accuracy: 52.78%

Name: Abdirahim Ahmed
Student ID: 801338570

Epoch 53/100, Loss: 0.6471, Accuracy: 52.77%
Epoch 54/100, Loss: 0.6358, Accuracy: 52.29%
Epoch 55/100, Loss: 0.6236, Accuracy: 52.35%
Epoch 56/100, Loss: 0.6134, Accuracy: 53.21%
Epoch 57/100, Loss: 0.6027, Accuracy: 51.03%
Epoch 58/100, Loss: 0.5928, Accuracy: 50.28%
Epoch 59/100, Loss: 0.5812, Accuracy: 52.22%
Epoch 60/100, Loss: 0.5720, Accuracy: 52.11%
Epoch 61/100, Loss: 0.5601, Accuracy: 52.02%
Epoch 62/100, Loss: 0.5519, Accuracy: 52.31%
Epoch 63/100, Loss: 0.5416, Accuracy: 52.68%
Epoch 64/100, Loss: 0.5309, Accuracy: 51.80%
Epoch 65/100, Loss: 0.5216, Accuracy: 50.98%
Epoch 66/100, Loss: 0.5122, Accuracy: 51.01%
Epoch 67/100, Loss: 0.5032, Accuracy: 51.49%
Epoch 68/100, Loss: 0.4938, Accuracy: 51.94%
Epoch 69/100, Loss: 0.4848, Accuracy: 51.62%
Epoch 70/100, Loss: 0.4751, Accuracy: 52.16%
Epoch 71/100, Loss: 0.4673, Accuracy: 51.03%
Epoch 72/100, Loss: 0.4590, Accuracy: 49.76%
Epoch 73/100, Loss: 0.4504, Accuracy: 51.44%
Epoch 74/100, Loss: 0.4409, Accuracy: 51.91%
Epoch 75/100, Loss: 0.4322, Accuracy: 50.60%
Epoch 76/100, Loss: 0.4242, Accuracy: 52.07%
Epoch 77/100, Loss: 0.4165, Accuracy: 52.05%
Epoch 78/100, Loss: 0.4074, Accuracy: 50.89%
Epoch 79/100, Loss: 0.3998, Accuracy: 51.58%
Epoch 80/100, Loss: 0.3933, Accuracy: 49.81%
Epoch 81/100, Loss: 0.3850, Accuracy: 51.60%
Epoch 82/100, Loss: 0.3756, Accuracy: 50.94%
Epoch 83/100, Loss: 0.3709, Accuracy: 51.35%
Epoch 84/100, Loss: 0.3636, Accuracy: 51.15%
Epoch 85/100, Loss: 0.3550, Accuracy: 51.10%
Epoch 86/100, Loss: 0.3487, Accuracy: 50.05%
Epoch 87/100, Loss: 0.3422, Accuracy: 50.79%
Epoch 88/100, Loss: 0.3348, Accuracy: 50.07%
Epoch 89/100, Loss: 0.3270, Accuracy: 51.70%
Epoch 90/100, Loss: 0.3227, Accuracy: 47.79%
Epoch 91/100, Loss: 0.3173, Accuracy: 48.67%
Epoch 92/100, Loss: 0.3090, Accuracy: 49.81%
Epoch 93/100, Loss: 0.3017, Accuracy: 50.04%
Epoch 94/100, Loss: 0.2970, Accuracy: 51.47%
Epoch 95/100, Loss: 0.2901, Accuracy: 50.37%

Name: Abdirahim Ahmed
Student ID: 801338570

Epoch 96/100, Loss: 0.2838, Accuracy: 51.31%
Epoch 97/100, Loss: 0.2775, Accuracy: 51.43%
Epoch 98/100, Loss: 0.2717, Accuracy: 51.28%
Epoch 99/100, Loss: 0.2663, Accuracy: 51.08%
Epoch 100/100, Loss: 0.2606, Accuracy: 50.72%
Problem 3a Results (Single Hidden Layer):
Training Time: 1804.400067806244 seconds
Train Losses (last epoch): 0.2606372979405286
Test Accuracy: 50.72 %

Problem 3a Analysis:
The single-hidden-layer model with 256 nodes achieved a final train loss of 0.2606 and a test accuracy of 50.72% after 100 epochs. The loss decreased steadily throughout training, while the test accuracy improved but plateaued around 50%. This is expected for a fully connected network on the CIFAR-10 dataset, as such models are less effective at handling image data compared to convolutional neural networks (CNNs). Despite the model's simplicity, it was able to classify the dataset with moderate accuracy, but it struggled to fully capture the spatial structure of image data.

b. Extend your network with two more additional hidden layers, like the example we did in lecture. Train your network for 100 epochs. Report your training time, loss, and evaluation accuracy after 100 epochs. Analyze your results in your report and compare your model size and accuracy over the baseline implementation in Problem1. a. Do you see any over-fitting? Make sure to submit your code by providing the GitHub URL of your course repository for this course. (35pt)
The network was extended to three hidden layers with 512, 256, and 128 nodes, respectively, to improve performance. While the train loss decreased significantly, the test accuracy improved only slightly (~54%), indicating overfitting. Fully connected layers were a limiting factor for handling the image data, and using CNNs would likely yield much better results

Epoch 1/100, Loss: 2.1856, Accuracy: 29.34%
Epoch 2/100, Loss: 1.8995, Accuracy: 36.61%
Epoch 3/100, Loss: 1.7320, Accuracy: 40.86%
Epoch 4/100, Loss: 1.6304, Accuracy: 43.67%
Epoch 5/100, Loss: 1.5553, Accuracy: 46.28%
Epoch 6/100, Loss: 1.4884, Accuracy: 47.91%
Epoch 7/100, Loss: 1.4295, Accuracy: 48.12%
Epoch 8/100, Loss: 1.3773, Accuracy: 49.80%
Epoch 9/100, Loss: 1.3295, Accuracy: 50.33%
Epoch 10/100, Loss: 1.2848, Accuracy: 51.81%
Epoch 11/100, Loss: 1.2439, Accuracy: 51.54%
Epoch 12/100, Loss: 1.2039, Accuracy: 51.77%

Name: Abdirahim Ahmed
Student ID: 801338570

Epoch 13/100, Loss: 1.1677, Accuracy: 52.24%
Epoch 14/100, Loss: 1.1329, Accuracy: 48.65%
Epoch 15/100, Loss: 1.0998, Accuracy: 51.94%
Epoch 16/100, Loss: 1.0651, Accuracy: 51.57%
Epoch 17/100, Loss: 1.0324, Accuracy: 52.92%
Epoch 18/100, Loss: 1.0002, Accuracy: 52.51%
Epoch 19/100, Loss: 0.9679, Accuracy: 50.49%
Epoch 20/100, Loss: 0.9386, Accuracy: 50.97%
Epoch 21/100, Loss: 0.9076, Accuracy: 52.44%
Epoch 22/100, Loss: 0.8751, Accuracy: 53.82%
Epoch 23/100, Loss: 0.8467, Accuracy: 50.12%
Epoch 24/100, Loss: 0.8166, Accuracy: 51.99%
Epoch 25/100, Loss: 0.7880, Accuracy: 51.71%
Epoch 26/100, Loss: 0.7596, Accuracy: 50.32%
Epoch 27/100, Loss: 0.7308, Accuracy: 52.69%
Epoch 28/100, Loss: 0.7058, Accuracy: 52.33%
Epoch 29/100, Loss: 0.6765, Accuracy: 47.74%
Epoch 30/100, Loss: 0.6476, Accuracy: 49.56%
Epoch 31/100, Loss: 0.6266, Accuracy: 53.46%
Epoch 32/100, Loss: 0.5942, Accuracy: 47.94%
Epoch 33/100, Loss: 0.5732, Accuracy: 47.40%
Epoch 34/100, Loss: 0.5437, Accuracy: 44.98%
Epoch 35/100, Loss: 0.5240, Accuracy: 48.46%
Epoch 36/100, Loss: 0.5023, Accuracy: 50.71%
Epoch 37/100, Loss: 0.4727, Accuracy: 45.92%
Epoch 38/100, Loss: 0.4592, Accuracy: 52.19%
Epoch 39/100, Loss: 0.4266, Accuracy: 49.34%
Epoch 40/100, Loss: 0.4094, Accuracy: 49.34%
Epoch 41/100, Loss: 0.3903, Accuracy: 51.16%
Epoch 42/100, Loss: 0.3699, Accuracy: 53.27%
Epoch 43/100, Loss: 0.3438, Accuracy: 53.54%
Epoch 44/100, Loss: 0.3367, Accuracy: 50.88%
Epoch 45/100, Loss: 0.3208, Accuracy: 51.36%
Epoch 46/100, Loss: 0.2939, Accuracy: 51.81%
Epoch 47/100, Loss: 0.2807, Accuracy: 51.75%
Epoch 48/100, Loss: 0.2688, Accuracy: 48.98%
Epoch 49/100, Loss: 0.2489, Accuracy: 52.09%
Epoch 50/100, Loss: 0.2391, Accuracy: 53.21%
Epoch 51/100, Loss: 0.2117, Accuracy: 51.85%
Epoch 52/100, Loss: 0.2090, Accuracy: 53.22%
Epoch 53/100, Loss: 0.2146, Accuracy: 51.24%
Epoch 54/100, Loss: 0.1924, Accuracy: 46.02%
Epoch 55/100, Loss: 0.1863, Accuracy: 51.41%

Name: Abdirahim Ahmed
Student ID: 801338570

Epoch 56/100, Loss: 0.1924, Accuracy: 47.01%
Epoch 57/100, Loss: 0.1564, Accuracy: 51.06%
Epoch 58/100, Loss: 0.1360, Accuracy: 51.98%
Epoch 59/100, Loss: 0.1248, Accuracy: 49.86%
Epoch 60/100, Loss: 0.1303, Accuracy: 52.50%
Epoch 61/100, Loss: 0.1207, Accuracy: 48.21%
Epoch 62/100, Loss: 0.1216, Accuracy: 53.78%
Epoch 63/100, Loss: 0.1129, Accuracy: 41.85%
Epoch 64/100, Loss: 0.1579, Accuracy: 48.51%
Epoch 65/100, Loss: 0.0713, Accuracy: 53.17%
Epoch 66/100, Loss: 0.0738, Accuracy: 52.66%
Epoch 67/100, Loss: 0.0688, Accuracy: 45.89%
Epoch 68/100, Loss: 0.1186, Accuracy: 53.64%
Epoch 69/100, Loss: 0.1106, Accuracy: 47.95%
Epoch 70/100, Loss: 0.1052, Accuracy: 51.31%
Epoch 71/100, Loss: 0.0622, Accuracy: 53.50%
Epoch 72/100, Loss: 0.0513, Accuracy: 51.24%
Epoch 73/100, Loss: 0.0514, Accuracy: 53.08%
Epoch 74/100, Loss: 0.0240, Accuracy: 53.94%
Epoch 75/100, Loss: 0.0173, Accuracy: 54.09%
Epoch 76/100, Loss: 0.0189, Accuracy: 52.88%
Epoch 77/100, Loss: 0.0143, Accuracy: 54.03%
Epoch 78/100, Loss: 0.0100, Accuracy: 52.82%
Epoch 79/100, Loss: 0.0098, Accuracy: 54.08%
Epoch 80/100, Loss: 0.0071, Accuracy: 54.22%
Epoch 81/100, Loss: 0.0057, Accuracy: 54.09%
Epoch 82/100, Loss: 0.0059, Accuracy: 50.91%
Epoch 83/100, Loss: 0.0075, Accuracy: 54.09%
Epoch 84/100, Loss: 0.0046, Accuracy: 54.21%
Epoch 85/100, Loss: 0.0044, Accuracy: 53.80%
Epoch 86/100, Loss: 0.0037, Accuracy: 54.22%
Epoch 87/100, Loss: 0.0035, Accuracy: 54.27%
Epoch 88/100, Loss: 0.0030, Accuracy: 54.27%
Epoch 89/100, Loss: 0.0030, Accuracy: 53.80%
Epoch 90/100, Loss: 0.0034, Accuracy: 54.40%
Epoch 91/100, Loss: 0.0026, Accuracy: 54.33%
Epoch 92/100, Loss: 0.0024, Accuracy: 54.31%
Epoch 93/100, Loss: 0.0024, Accuracy: 54.11%
Epoch 94/100, Loss: 0.0024, Accuracy: 54.18%
Epoch 95/100, Loss: 0.0024, Accuracy: 54.23%
Epoch 96/100, Loss: 0.0020, Accuracy: 54.14%
Epoch 97/100, Loss: 0.0020, Accuracy: 54.16%
Epoch 98/100, Loss: 0.0019, Accuracy: 54.10%

Name: Abdirahim Ahmed
Student ID: 801338570

Epoch 99/100, Loss: 0.0018, Accuracy: 53.95%
Epoch 100/100, Loss: 0.0018, Accuracy: 53.94%
Problem 3b Results (Three Hidden Layers):
Training Time: 2270.0692205429077 seconds
Train Losses (last epoch): 0.0017761932265417189
Test Accuracy: 53.94 %

Problem 3b Analysis:
The three-hidden-layer model with 512, 256, and 128 nodes achieved a final train loss of 0.0018 and a test accuracy of 53.94% after 100 epochs. While the train loss decreased to an extremely low value, the test accuracy improved only slightly compared to the single-hidden-layer model (50.72% vs. 53.94%). The additional layers increased the network’s capacity, which helped achieve a lower train loss but also led to overfitting. This is evident from the significant gap between train loss and test accuracy. The use of fully connected layers further limited the model's ability to handle the image data effectively.