

AP_Project

December 22, 2023

1 Gradient Accumulation Project

The following blocks of code are part of a research project investigating the effectiveness of using gradient accumulation to enhance neural network learning. With this code my aim is to show that taking one “big” step across the entire data set may prove to be beneficial in several specific contexts. My goal is to create two identical models and only compare the different training approaches and results. It is important to note that the code below only serves as the exploratory first steps of this research project. It is also worth noting, that while PyTorch already provides this functionality, Tensorflow & Keras do not. Therefore, this implementation may possibly be integrated into a future version of the Keras library.

1.1 Import Relevant Packages

These packages can either be pre-installed directly by the user, or the user can opt to use the poetry package to install the project’s dependencies. To use poetry, first run: “poetry install –no-root”, then run “poetry shell” to activate the virtual environment in which to run this notebook.

```
[ ]: import tensorflow as tf
import numpy as np
from copy import deepcopy
import matplotlib.pyplot as plt
#from micrograd.engine import Value
from tensorflow import keras as keras
from keras.losses import SparseCategoricalCrossentropy
from keras.optimizers import Adam
from keras.models import Sequential
from keras.layers import Dense
```

If you are running this notebook on an Apple silicon chip (M1/M2), please make sure that TensorFlow has GPU access (not absolutely necessary, but it does improve training performance).

```
[ ]: # Check for TensorFlow GPU access
print(f"TensorFlow has access to the following devices:\n{tf.config.
↵list_physical_devices()}")
```

```
TensorFlow has access to the following devices:
[PhysicalDevice(name='/physical_device:CPU:0', device_type='CPU'),
PhysicalDevice(name='/physical_device:GPU:0', device_type='GPU')]
```

2 Load the Data

Now, the first step is to load our data. For simplicity, we will be using data sets that Keras provides us, specifically MNIST & Fashion-MNIST.

```
[ ]: # Decide what batch size you want to train the models with
batch_size = 512

# Load and preprocess the standard MNIST data
mnist = keras.datasets.mnist
(X_train_mnist, y_train_mnist), (X_test_mnist, y_test_mnist) = mnist.load_data()
X_train_mnist = X_train_mnist / 255.0
X_test_mnist = X_test_mnist / 255.0

# Create the MNIST dataset
data_mnist = tf.data.Dataset.from_tensor_slices((X_train_mnist, y_train_mnist)).
    ↪ batch(batch_size)

# Load and preprocess the Fashion MNIST data
fashion_mnist = keras.datasets.fashion_mnist
(X_train_fashion, y_train_fashion), (X_test_fashion, y_test_fashion) = ↪
    ↪ fashion_mnist.load_data()
X_train_fashion = X_train_fashion / 255.0
X_test_fashion = X_test_fashion / 255.0

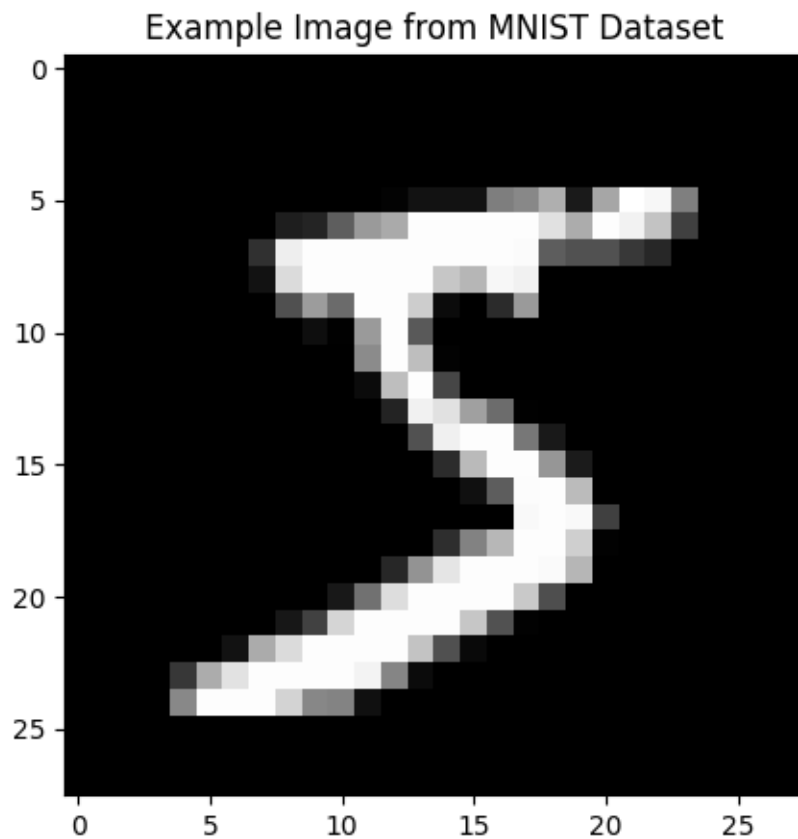
# Create the Fashion MNIST dataset
data_fashion = tf.data.Dataset.from_tensor_slices((X_train_fashion, ↪
    ↪ y_train_fashion)).batch(batch_size)
```

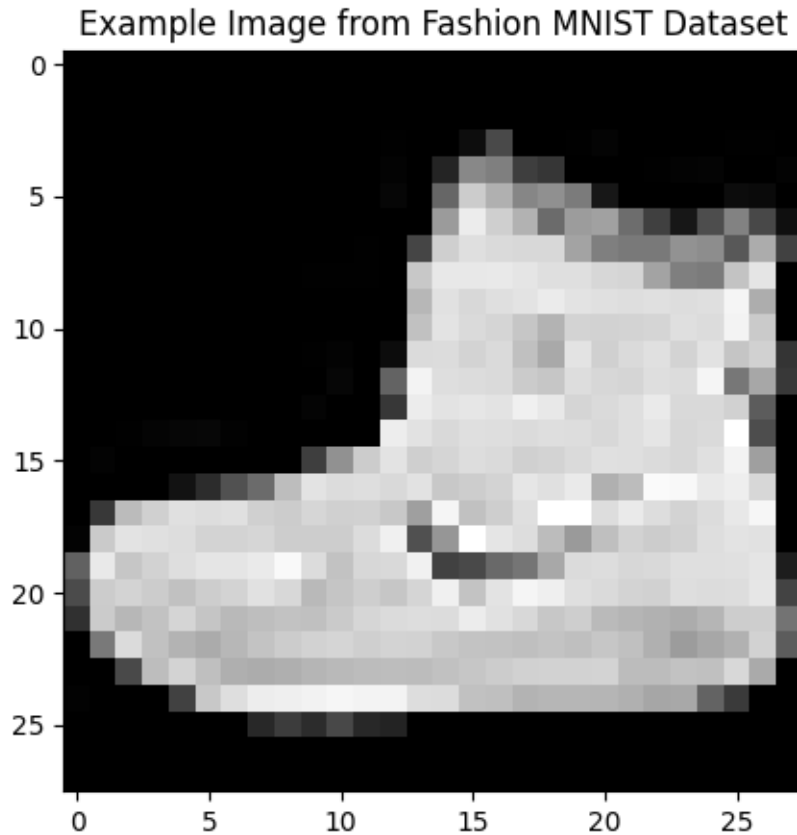
```
2023-12-22 12:07:44.691296: I metal_plugin/src/device/metal_device.cc:1154]
Metal device set to: Apple M2 Pro
2023-12-22 12:07:44.691319: I metal_plugin/src/device/metal_device.cc:296]
systemMemory: 16.00 GB
2023-12-22 12:07:44.691329: I metal_plugin/src/device/metal_device.cc:313]
maxCacheSize: 5.33 GB
2023-12-22 12:07:44.691543: I
tensorflow/core/common_runtime/pluggable_device/pluggable_device_factory.cc:306]
Could not identify NUMA node of platform GPU ID 0, defaulting to 0. Your kernel
may not have been built with NUMA support.
2023-12-22 12:07:44.692349: I
tensorflow/core/common_runtime/pluggable_device/pluggable_device_factory.cc:272]
Created TensorFlow device (/job:localhost/replica:0/task:0/device:GPU:0 with 0
MB memory) -> physical PluggableDevice (device: 0, name: METAL, pci bus id:
<undefined>)
```

Let's take a look at an example image of both of our data sets. To do so, we will need to visualize the arrays using matplotlib.

```
[ ]: # Plot an example image from the MNIST dataset
plt.figure(figsize=(5,5))
plt.imshow(X_train_mnist[0], cmap='gray')
plt.title('Example Image from MNIST Dataset')
plt.show()

# Plot an example image from the Fashion MNIST dataset
plt.figure(figsize=(5,5))
plt.imshow(X_train_fashion[0], cmap='gray')
plt.title('Example Image from Fashion MNIST Dataset')
plt.show()
```





It looks like the first image should be the digit “5”, although it looks very much like an “S”. The second image seems to be a shoe, specifically a boot. Please note that we would normally explore out data set further to get to know the intricacies of it. However, since these are two of the most commonly used data sets, I will not spend more time exploring them here.

3 Create the Models

Here we create the two models we will be using. In order to make sure they have the same initial weights, we need to set a seed. For the model, I have decided to use 3 Dense layers with the number of neurons descending in half steps, until finally arriving at the output/prediction layer with 10 neurons, representing the number of classes in both data sets. This is common practice, since deeper layers will learn more abstract patterns, after the first layer learns more “holistically”. Also since the size of our training set is 60’000, we ideally would like to have between 60’000-150’000 parameters in our model - 1/4-1/10 of total possible option space (600’000). Here our model has exactly 111’146 parameters.

```
[ ]: # Define a function to create the model
def create_model():
    model = keras.Sequential([
        keras.layers.Flatten(input_shape=(28, 28)),
```

```

        keras.layers.Dense(128, activation='tanh'),
        keras.layers.Dense(64, activation='tanh'),
        keras.layers.Dense(32, activation='tanh'),
        keras.layers.Dense(10, activation='linear')
    ])
    return model

# Set the seed
tf.random.set_seed(42)

# Create the standard model
s_model = create_model()

# Create the accumulation model
a_model = create_model()

s_model.summary()

```

Model: "sequential_14"

Layer (type)	Output Shape	Param #
flatten_14 (Flatten)	(None, 784)	0
dense_56 (Dense)	(None, 128)	100480
dense_57 (Dense)	(None, 64)	8256
dense_58 (Dense)	(None, 32)	2080
dense_59 (Dense)	(None, 10)	330

```

Total params: 111146 (434.16 KB)
Trainable params: 111146 (434.16 KB)
Non-trainable params: 0 (0.00 Byte)

```

Layer (type)	Output Shape	Param #
flatten_14 (Flatten)	(None, 784)	0
dense_56 (Dense)	(None, 128)	100480
dense_57 (Dense)	(None, 64)	8256
dense_58 (Dense)	(None, 32)	2080

dense_59 (Dense) (None, 10) 330

```
=====
Total params: 111146 (434.16 KB)
Trainable params: 111146 (434.16 KB)
Non-trainable params: 0 (0.00 Byte)
-----
```

Next we define the loss function and the optimizer. We use the Adam optimizer with a learning rate of 0.01 and the sparse categorical crossentropy loss function. Please note that for the purposes of this research project, the model and its optimizers, loss functions, activation functions etc. are irrelevant, as we are only interested in comparing the training approaches with each other. As long as both models are identical, we can draw a fair comparison.

```
[ ]: # Define the optimizer
optimizer = tf.keras.optimizers.legacy.SGD(learning_rate = 0.01)

# Explicitly define loss function with (logits = True) because we are not using
↳ softmax activation in the final layer
loss_fn = keras.losses.SparseCategoricalCrossentropy(from_logits=True)

s_model.compile(optimizer=optimizer, loss=loss_fn, metrics=['accuracy'])

a_model.compile(optimizer=optimizer, loss=loss_fn, metrics=['accuracy'])
```

4 Train Standard Model

Now we move on the training. Since we want to do this step more or less completely manually, while still tracking relevant metrics, it requires a lengthy set-up. We first create the variables that will track our metrics. We then decide on how many epochs we want to train for. Feel free to adjust this based on how long you wish to train the model on your local machine. The default it is set to is 200 epochs as that gives us a good approximation.

```
[ ]: # Initialize lists to store metrics
train_losses_standard = []
val_losses_standard = []
train_accuracies_standard = []
val_accuracies_standard = []

# Number of epochs to train for
num_epochs = 200

for epoch in range(num_epochs):
    print(f"Epoch {epoch+1}/{num_epochs}")

    # Initialize lists to store metrics for each batch
    batch_losses = []
    batch_accuracies = []
```

```

# Iterate over the batches of the dataset
for step, (X_batch, y_batch) in enumerate(data_mnist):
    with tf.GradientTape() as tape:
        # Forward pass
        y_pred = s_model(X_batch, training=True)
        # Compute the loss value
        loss_value = loss_fn(y_batch, y_pred)

    # Compute gradients
    grads = tape.gradient(loss_value, s_model.trainable_weights)

    # Update weights
    optimizer.apply_gradients(zip(grads, s_model.trainable_weights))

    # Store metrics for this batch
    batch_losses.append(loss_value.numpy())
    batch_accuracies.append(tf.reduce_mean(tf.cast(tf.equal(tf.
↪ cast(y_batch, tf.int64), tf.argmax(y_pred, axis=1)), tf.float32)).numpy()))

    # Compute average metrics for this epoch
    train_loss = np.mean(batch_losses)
    train_accuracy = np.mean(batch_accuracies)
    train_losses_standard.append(train_loss)
    train_accuracies_standard.append(train_accuracy)

    # Compute validation metrics for this epoch
    y_val_pred = s_model(X_test_mnist, training=False)
    val_loss = np.mean(loss_fn(y_test_mnist, y_val_pred).numpy())
    val_accuracy = np.mean(tf.cast(tf.equal(y_test_mnist, tf.argmax(y_val_pred,
↪ axis=1)), tf.float32).numpy())
    val_losses_standard.append(val_loss)
    val_accuracies_standard.append(val_accuracy)

    # Log metrics at the end of each epoch
    print(f"Training loss for epoch {epoch+1}: {train_loss}")
    print(f"Training accuracy for epoch {epoch+1}: {train_accuracy}")
    print(f"Validation loss for epoch {epoch+1}: {val_loss}")
    print(f"Validation accuracy for epoch {epoch+1}: {val_accuracy}")

```

Epoch 1/200

Training loss for epoch 1: 1.639849305152893

Training accuracy for epoch 1: 0.5337713956832886

Validation loss for epoch 1: 1.2186195850372314

Validation accuracy for epoch 1: 0.7346000075340271

Epoch 2/200
Training loss for epoch 2: 1.0583500862121582
Training accuracy for epoch 2: 0.7658015489578247
Validation loss for epoch 2: 0.9000820517539978
Validation accuracy for epoch 2: 0.8077999949455261
Epoch 3/200
Training loss for epoch 3: 0.8312516212463379
Training accuracy for epoch 3: 0.8174269795417786
Validation loss for epoch 3: 0.7373012900352478
Validation accuracy for epoch 3: 0.8420000076293945
Epoch 4/200
Training loss for epoch 4: 0.7018334865570068
Training accuracy for epoch 4: 0.8442245125770569
Validation loss for epoch 4: 0.6351539492607117
Validation accuracy for epoch 4: 0.8636000156402588
Epoch 5/200
Training loss for epoch 5: 0.616486132144928
Training accuracy for epoch 5: 0.8611571192741394
Validation loss for epoch 5: 0.5647680759429932
Validation accuracy for epoch 5: 0.8772000074386597
Epoch 6/200
Training loss for epoch 6: 0.5559747219085693
Training accuracy for epoch 6: 0.8722744584083557
Validation loss for epoch 6: 0.5136528611183167
Validation accuracy for epoch 6: 0.8842999935150146
Epoch 7/200
Training loss for epoch 7: 0.51109778881073
Training accuracy for epoch 7: 0.8800538182258606
Validation loss for epoch 7: 0.47516053915023804
Validation accuracy for epoch 7: 0.8907999992370605
Epoch 8/200
Training loss for epoch 8: 0.4766659438610077
Training accuracy for epoch 8: 0.8856649398803711
Validation loss for epoch 8: 0.44529959559440613
Validation accuracy for epoch 8: 0.8942999839782715
Epoch 9/200
Training loss for epoch 9: 0.4494772255420685
Training accuracy for epoch 9: 0.889676034450531
Validation loss for epoch 9: 0.42151758074760437
Validation accuracy for epoch 9: 0.8985000252723694
Epoch 10/200
Training loss for epoch 10: 0.42745712399482727
Training accuracy for epoch 10: 0.893400251865387
Validation loss for epoch 10: 0.40212488174438477
Validation accuracy for epoch 10: 0.9016000032424927
Epoch 11/200
Training loss for epoch 11: 0.4092199504375458
Training accuracy for epoch 11: 0.8965560793876648

Validation loss for epoch 11: 0.3859759569168091
Validation accuracy for epoch 11: 0.9039000272750854
Epoch 12/200
Training loss for epoch 12: 0.39381688833236694
Training accuracy for epoch 12: 0.8986912965774536
Validation loss for epoch 12: 0.37227728962898254
Validation accuracy for epoch 12: 0.9057000279426575
Epoch 13/200
Training loss for epoch 13: 0.3805834949016571
Training accuracy for epoch 13: 0.901339590549469
Validation loss for epoch 13: 0.3604678809642792
Validation accuracy for epoch 13: 0.9085000157356262
Epoch 14/200
Training loss for epoch 14: 0.3690442442893982
Training accuracy for epoch 14: 0.9032155275344849
Validation loss for epoch 14: 0.35014310479164124
Validation accuracy for epoch 14: 0.9097999930381775
Epoch 15/200
Training loss for epoch 15: 0.3588520288467407
Training accuracy for epoch 15: 0.9047603011131287
Validation loss for epoch 15: 0.34100526571273804
Validation accuracy for epoch 15: 0.9111999869346619
Epoch 16/200
Training loss for epoch 16: 0.3497481644153595
Training accuracy for epoch 16: 0.9064486026763916
Validation loss for epoch 16: 0.3328314423561096
Validation accuracy for epoch 16: 0.9126999974250793
Epoch 17/200
Training loss for epoch 17: 0.3415367007255554
Training accuracy for epoch 17: 0.9079548716545105
Validation loss for epoch 17: 0.3254516124725342
Validation accuracy for epoch 17: 0.9135000109672546
Epoch 18/200
Training loss for epoch 18: 0.3340665102005005
Training accuracy for epoch 18: 0.9096651673316956
Validation loss for epoch 18: 0.3187342882156372
Validation accuracy for epoch 18: 0.9143000245094299
Epoch 19/200
Training loss for epoch 19: 0.32721948623657227
Training accuracy for epoch 19: 0.9108072519302368
Validation loss for epoch 19: 0.3125758171081543
Validation accuracy for epoch 19: 0.9161999821662903
Epoch 20/200
Training loss for epoch 20: 0.32090190052986145
Training accuracy for epoch 20: 0.9124459028244019
Validation loss for epoch 20: 0.3068937361240387
Validation accuracy for epoch 20: 0.9172000288963318
Epoch 21/200

Training loss for epoch 21: 0.3150383532047272
Training accuracy for epoch 21: 0.9135714173316956
Validation loss for epoch 21: 0.3016214966773987
Validation accuracy for epoch 21: 0.9182000160217285
Epoch 22/200
Training loss for epoch 22: 0.3095676600933075
Training accuracy for epoch 22: 0.9147631525993347
Validation loss for epoch 22: 0.29670456051826477
Validation accuracy for epoch 22: 0.9192000031471252
Epoch 23/200
Training loss for epoch 23: 0.3044395446777344
Training accuracy for epoch 23: 0.9157397150993347
Validation loss for epoch 23: 0.2920980751514435
Validation accuracy for epoch 23: 0.9196000099182129
Epoch 24/200
Training loss for epoch 24: 0.2996121942996979
Training accuracy for epoch 24: 0.9167162775993347
Validation loss for epoch 24: 0.2877645790576935
Validation accuracy for epoch 24: 0.9205999970436096
Epoch 25/200
Training loss for epoch 25: 0.2950507700443268
Training accuracy for epoch 25: 0.917941153049469
Validation loss for epoch 25: 0.2836725115776062
Validation accuracy for epoch 25: 0.9223999977111816
Epoch 26/200
Training loss for epoch 26: 0.29072579741477966
Training accuracy for epoch 26: 0.9189011454582214
Validation loss for epoch 26: 0.27979516983032227
Validation accuracy for epoch 26: 0.9236000180244446
Epoch 27/200
Training loss for epoch 27: 0.2866121828556061
Training accuracy for epoch 27: 0.9195687770843506
Validation loss for epoch 27: 0.2761096954345703
Validation accuracy for epoch 27: 0.9247000217437744
Epoch 28/200
Training loss for epoch 28: 0.2826884388923645
Training accuracy for epoch 28: 0.9207274317741394
Validation loss for epoch 28: 0.27259644865989685
Validation accuracy for epoch 28: 0.9258000254631042
Epoch 29/200
Training loss for epoch 29: 0.2789359986782074
Training accuracy for epoch 29: 0.9215550422668457
Validation loss for epoch 29: 0.2692383825778961
Validation accuracy for epoch 29: 0.9269999861717224
Epoch 30/200
Training loss for epoch 30: 0.2753389775753021
Training accuracy for epoch 30: 0.9226198196411133
Validation loss for epoch 30: 0.2660207152366638

Validation accuracy for epoch 30: 0.9280999898910522
Epoch 31/200
Training loss for epoch 31: 0.2718832492828369
Training accuracy for epoch 31: 0.9234970808029175
Validation loss for epoch 31: 0.26293057203292847
Validation accuracy for epoch 31: 0.9290000200271606
Epoch 32/200
Training loss for epoch 32: 0.2685565948486328
Training accuracy for epoch 32: 0.9242474436759949
Validation loss for epoch 32: 0.2599565088748932
Validation accuracy for epoch 32: 0.9294000267982483
Epoch 33/200
Training loss for epoch 33: 0.26534831523895264
Training accuracy for epoch 33: 0.9249094724655151
Validation loss for epoch 33: 0.2570885717868805
Validation accuracy for epoch 33: 0.9301999807357788
Epoch 34/200
Training loss for epoch 34: 0.2622488737106323
Training accuracy for epoch 34: 0.9258198738098145
Validation loss for epoch 34: 0.25431790947914124
Validation accuracy for epoch 34: 0.9308000206947327
Epoch 35/200
Training loss for epoch 35: 0.2592497766017914
Training accuracy for epoch 35: 0.9267578125
Validation loss for epoch 35: 0.2516367733478546
Validation accuracy for epoch 35: 0.9312000274658203
Epoch 36/200
Training loss for epoch 36: 0.2563436031341553
Training accuracy for epoch 36: 0.9277674555778503
Validation loss for epoch 36: 0.24903813004493713
Validation accuracy for epoch 36: 0.9312999844551086
Epoch 37/200
Training loss for epoch 37: 0.25352364778518677
Training accuracy for epoch 37: 0.9286778569221497
Validation loss for epoch 37: 0.24651582539081573
Validation accuracy for epoch 37: 0.9319000244140625
Epoch 38/200
Training loss for epoch 38: 0.25078392028808594
Training accuracy for epoch 38: 0.9292240738868713
Validation loss for epoch 38: 0.2440643608570099
Validation accuracy for epoch 38: 0.9326000213623047
Epoch 39/200
Training loss for epoch 39: 0.2481190264225006
Training accuracy for epoch 39: 0.930084764957428
Validation loss for epoch 39: 0.24167875945568085
Validation accuracy for epoch 39: 0.9329000115394592
Epoch 40/200
Training loss for epoch 40: 0.24552419781684875

Training accuracy for epoch 40: 0.9306971430778503
Validation loss for epoch 40: 0.2393546849489212
Validation accuracy for epoch 40: 0.9333999752998352
Epoch 41/200
Training loss for epoch 41: 0.24299505352973938
Training accuracy for epoch 41: 0.931326150894165
Validation loss for epoch 41: 0.2370881885290146
Validation accuracy for epoch 41: 0.9337999820709229
Epoch 42/200
Training loss for epoch 42: 0.2405276745557785
Training accuracy for epoch 42: 0.931955099105835
Validation loss for epoch 42: 0.23487573862075806
Validation accuracy for epoch 42: 0.9344000220298767
Epoch 43/200
Training loss for epoch 43: 0.2381184995174408
Training accuracy for epoch 43: 0.9325178861618042
Validation loss for epoch 43: 0.23271413147449493
Validation accuracy for epoch 43: 0.9347000122070312
Epoch 44/200
Training loss for epoch 44: 0.2357642948627472
Training accuracy for epoch 44: 0.9333785772323608
Validation loss for epoch 44: 0.2306005358695984
Validation accuracy for epoch 44: 0.9351999759674072
Epoch 45/200
Training loss for epoch 45: 0.2334621250629425
Training accuracy for epoch 45: 0.9339744448661804
Validation loss for epoch 45: 0.22853241860866547
Validation accuracy for epoch 45: 0.9355999827384949
Epoch 46/200
Training loss for epoch 46: 0.2312094122171402
Training accuracy for epoch 46: 0.9346696138381958
Validation loss for epoch 46: 0.22650742530822754
Validation accuracy for epoch 46: 0.9365000128746033
Epoch 47/200
Training loss for epoch 47: 0.22900360822677612
Training accuracy for epoch 47: 0.9351165294647217
Validation loss for epoch 47: 0.22452352941036224
Validation accuracy for epoch 47: 0.9369000196456909
Epoch 48/200
Training loss for epoch 48: 0.22684256732463837
Training accuracy for epoch 48: 0.9359606504440308
Validation loss for epoch 48: 0.22257885336875916
Validation accuracy for epoch 48: 0.9373000264167786
Epoch 49/200
Training loss for epoch 49: 0.22472423315048218
Training accuracy for epoch 49: 0.9364737868309021
Validation loss for epoch 49: 0.22067169845104218
Validation accuracy for epoch 49: 0.9377999901771545

Epoch 50/200
Training loss for epoch 50: 0.22264674305915833
Training accuracy for epoch 50: 0.937075138092041
Validation loss for epoch 50: 0.21880052983760834
Validation accuracy for epoch 50: 0.9380999803543091
Epoch 51/200
Training loss for epoch 51: 0.2206083983182907
Training accuracy for epoch 51: 0.9375551342964172
Validation loss for epoch 51: 0.2169639617204666
Validation accuracy for epoch 51: 0.9381999969482422
Epoch 52/200
Training loss for epoch 52: 0.2186075896024704
Training accuracy for epoch 52: 0.9380186200141907
Validation loss for epoch 52: 0.21516074240207672
Validation accuracy for epoch 52: 0.9383000135421753
Epoch 53/200
Training loss for epoch 53: 0.21664293110370636
Training accuracy for epoch 53: 0.9385316967964172
Validation loss for epoch 53: 0.21338973939418793
Validation accuracy for epoch 53: 0.9388999938964844
Epoch 54/200
Training loss for epoch 54: 0.2147129476070404
Training accuracy for epoch 54: 0.939028263092041
Validation loss for epoch 54: 0.21164995431900024
Validation accuracy for epoch 54: 0.9391000270843506
Epoch 55/200
Training loss for epoch 55: 0.21281655132770538
Training accuracy for epoch 55: 0.9395579099655151
Validation loss for epoch 55: 0.20994040369987488
Validation accuracy for epoch 55: 0.939300000667572
Epoch 56/200
Training loss for epoch 56: 0.2109525054693222
Training accuracy for epoch 56: 0.9401869177818298
Validation loss for epoch 56: 0.20826022326946259
Validation accuracy for epoch 56: 0.9401999711990356
Epoch 57/200
Training loss for epoch 57: 0.2091197520494461
Training accuracy for epoch 57: 0.9407992959022522
Validation loss for epoch 57: 0.20660866796970367
Validation accuracy for epoch 57: 0.9405999779701233
Epoch 58/200
Training loss for epoch 58: 0.20731724798679352
Training accuracy for epoch 58: 0.9413014650344849
Validation loss for epoch 58: 0.20498493313789368
Validation accuracy for epoch 58: 0.9406999945640564
Epoch 59/200
Training loss for epoch 59: 0.2055441439151764
Training accuracy for epoch 59: 0.9415827989578247

Validation loss for epoch 59: 0.20338842272758484
Validation accuracy for epoch 59: 0.941100001335144
Epoch 60/200
Training loss for epoch 60: 0.2037995457649231
Training accuracy for epoch 60: 0.942228376865387
Validation loss for epoch 60: 0.20181845128536224
Validation accuracy for epoch 60: 0.9413999915122986
Epoch 61/200
Training loss for epoch 61: 0.2020825892686844
Training accuracy for epoch 61: 0.9425593614578247
Validation loss for epoch 61: 0.20027446746826172
Validation accuracy for epoch 61: 0.9417999982833862
Epoch 62/200
Training loss for epoch 62: 0.20039251446723938
Training accuracy for epoch 62: 0.9429566264152527
Validation loss for epoch 62: 0.1987559199333191
Validation accuracy for epoch 62: 0.9423999786376953
Epoch 63/200
Training loss for epoch 63: 0.19872859120368958
Training accuracy for epoch 63: 0.9433704018592834
Validation loss for epoch 63: 0.19726233184337616
Validation accuracy for epoch 63: 0.9426000118255615
Epoch 64/200
Training loss for epoch 64: 0.19709019362926483
Training accuracy for epoch 64: 0.9439994096755981
Validation loss for epoch 64: 0.19579312205314636
Validation accuracy for epoch 64: 0.9430000185966492
Epoch 65/200
Training loss for epoch 65: 0.19547659158706665
Training accuracy for epoch 65: 0.9443469643592834
Validation loss for epoch 65: 0.19434790313243866
Validation accuracy for epoch 65: 0.9429000020027161
Epoch 66/200
Training loss for epoch 66: 0.19388726353645325
Training accuracy for epoch 66: 0.9447938799858093
Validation loss for epoch 66: 0.19292622804641724
Validation accuracy for epoch 66: 0.9430999755859375
Epoch 67/200
Training loss for epoch 67: 0.1923215091228485
Training accuracy for epoch 67: 0.9451414942741394
Validation loss for epoch 67: 0.1915276199579239
Validation accuracy for epoch 67: 0.9437000155448914
Epoch 68/200
Training loss for epoch 68: 0.19077885150909424
Training accuracy for epoch 68: 0.9456877112388611
Validation loss for epoch 68: 0.19015173614025116
Validation accuracy for epoch 68: 0.9445000290870667
Epoch 69/200

Training loss for epoch 69: 0.18925870954990387
Training accuracy for epoch 69: 0.9461842775344849
Validation loss for epoch 69: 0.18879809975624084
Validation accuracy for epoch 69: 0.9448999762535095
Epoch 70/200
Training loss for epoch 70: 0.187760591506958
Training accuracy for epoch 70: 0.9465484023094177
Validation loss for epoch 70: 0.1874663531780243
Validation accuracy for epoch 70: 0.9452000260353088
Epoch 71/200
Training loss for epoch 71: 0.18628402054309845
Training accuracy for epoch 71: 0.9469290971755981
Validation loss for epoch 71: 0.18615610897541046
Validation accuracy for epoch 71: 0.9456999897956848
Epoch 72/200
Training loss for epoch 72: 0.1848284900188446
Training accuracy for epoch 72: 0.9472601413726807
Validation loss for epoch 72: 0.18486694991588593
Validation accuracy for epoch 72: 0.9459999799728394
Epoch 73/200
Training loss for epoch 73: 0.18339353799819946
Training accuracy for epoch 73: 0.9476739168167114
Validation loss for epoch 73: 0.1835985779762268
Validation accuracy for epoch 73: 0.9462000131607056
Epoch 74/200
Training loss for epoch 74: 0.18197883665561676
Training accuracy for epoch 74: 0.9480215311050415
Validation loss for epoch 74: 0.18235056102275848
Validation accuracy for epoch 74: 0.9466999769210815
Epoch 75/200
Training loss for epoch 75: 0.18058384954929352
Training accuracy for epoch 75: 0.9486339092254639
Validation loss for epoch 75: 0.1811225414276123
Validation accuracy for epoch 75: 0.9470000267028809
Epoch 76/200
Training loss for epoch 76: 0.1792081892490387
Training accuracy for epoch 76: 0.949246346950531
Validation loss for epoch 76: 0.17991416156291962
Validation accuracy for epoch 76: 0.9477999806404114
Epoch 77/200
Training loss for epoch 77: 0.17785145342350006
Training accuracy for epoch 77: 0.949560821056366
Validation loss for epoch 77: 0.17872509360313416
Validation accuracy for epoch 77: 0.948199987411499
Epoch 78/200
Training loss for epoch 78: 0.17651331424713135
Training accuracy for epoch 78: 0.9501401782035828
Validation loss for epoch 78: 0.1775549203157425

Validation accuracy for epoch 78: 0.9487000107765198
Epoch 79/200
Training loss for epoch 79: 0.17519336938858032
Training accuracy for epoch 79: 0.9506036043167114
Validation loss for epoch 79: 0.17640331387519836
Validation accuracy for epoch 79: 0.9490000009536743
Epoch 80/200
Training loss for epoch 80: 0.17389120161533356
Training accuracy for epoch 80: 0.9509180784225464
Validation loss for epoch 80: 0.1752699464559555
Validation accuracy for epoch 80: 0.9492999911308289
Epoch 81/200
Training loss for epoch 81: 0.17260651290416718
Training accuracy for epoch 81: 0.9513319134712219
Validation loss for epoch 81: 0.17415441572666168
Validation accuracy for epoch 81: 0.949400007724762
Epoch 82/200
Training loss for epoch 82: 0.17133891582489014
Training accuracy for epoch 82: 0.9518284797668457
Validation loss for epoch 82: 0.17305640876293182
Validation accuracy for epoch 82: 0.9498999714851379
Epoch 83/200
Training loss for epoch 83: 0.17008812725543976
Training accuracy for epoch 83: 0.952176034450531
Validation loss for epoch 83: 0.17197559773921967
Validation accuracy for epoch 83: 0.9498999714851379
Epoch 84/200
Training loss for epoch 84: 0.1688537448644638
Training accuracy for epoch 84: 0.9526891708374023
Validation loss for epoch 84: 0.17091159522533417
Validation accuracy for epoch 84: 0.9502999782562256
Epoch 85/200
Training loss for epoch 85: 0.16763551533222198
Training accuracy for epoch 85: 0.952838122844696
Validation loss for epoch 85: 0.16986411809921265
Validation accuracy for epoch 85: 0.9508000016212463
Epoch 86/200
Training loss for epoch 86: 0.16643309593200684
Training accuracy for epoch 86: 0.9530698657035828
Validation loss for epoch 86: 0.16883276402950287
Validation accuracy for epoch 86: 0.9508000016212463
Epoch 87/200
Training loss for epoch 87: 0.16524620354175568
Training accuracy for epoch 87: 0.9534339904785156
Validation loss for epoch 87: 0.16781726479530334
Validation accuracy for epoch 87: 0.9510999917984009
Epoch 88/200
Training loss for epoch 88: 0.16407449543476105

Training accuracy for epoch 88: 0.9537319540977478
Validation loss for epoch 88: 0.16681727766990662
Validation accuracy for epoch 88: 0.95169997215271
Epoch 89/200
Training loss for epoch 89: 0.16291765868663788
Training accuracy for epoch 89: 0.9540795087814331
Validation loss for epoch 89: 0.16583245992660522
Validation accuracy for epoch 89: 0.9520999789237976
Epoch 90/200
Training loss for epoch 90: 0.16177543997764587
Training accuracy for epoch 90: 0.9542450308799744
Validation loss for epoch 90: 0.1648625284433365
Validation accuracy for epoch 90: 0.9523000121116638
Epoch 91/200
Training loss for epoch 91: 0.16064758598804474
Training accuracy for epoch 91: 0.9545429944992065
Validation loss for epoch 91: 0.16390712559223175
Validation accuracy for epoch 91: 0.9524999856948853
Epoch 92/200
Training loss for epoch 92: 0.1595337986946106
Training accuracy for epoch 92: 0.9549567699432373
Validation loss for epoch 92: 0.16296601295471191
Validation accuracy for epoch 92: 0.9528999924659729
Epoch 93/200
Training loss for epoch 93: 0.15843376517295837
Training accuracy for epoch 93: 0.9553043842315674
Validation loss for epoch 93: 0.16203884780406952
Validation accuracy for epoch 93: 0.9531000256538391
Epoch 94/200
Training loss for epoch 94: 0.15734727680683136
Training accuracy for epoch 94: 0.9556353688240051
Validation loss for epoch 94: 0.1611253172159195
Validation accuracy for epoch 94: 0.9535999894142151
Epoch 95/200
Training loss for epoch 95: 0.15627402067184448
Training accuracy for epoch 95: 0.9558008909225464
Validation loss for epoch 95: 0.16022516787052155
Validation accuracy for epoch 95: 0.9535999894142151
Epoch 96/200
Training loss for epoch 96: 0.15521380305290222
Training accuracy for epoch 96: 0.9561319351196289
Validation loss for epoch 96: 0.159338116645813
Validation accuracy for epoch 96: 0.9535999894142151
Epoch 97/200
Training loss for epoch 97: 0.1541663557291031
Training accuracy for epoch 97: 0.9564629793167114
Validation loss for epoch 97: 0.15846383571624756
Validation accuracy for epoch 97: 0.9538000226020813

Epoch 98/200
Training loss for epoch 98: 0.1531313806772232
Training accuracy for epoch 98: 0.9566947221755981
Validation loss for epoch 98: 0.15760213136672974
Validation accuracy for epoch 98: 0.9539999961853027
Epoch 99/200
Training loss for epoch 99: 0.15210872888565063
Training accuracy for epoch 99: 0.9569098949432373
Validation loss for epoch 99: 0.15675266087055206
Validation accuracy for epoch 99: 0.9541000127792358
Epoch 100/200
Training loss for epoch 100: 0.1510981023311615
Training accuracy for epoch 100: 0.957141637802124
Validation loss for epoch 100: 0.15591521561145782
Validation accuracy for epoch 100: 0.9542999863624573
Epoch 101/200
Training loss for epoch 101: 0.15009929239749908
Training accuracy for epoch 101: 0.9574726819992065
Validation loss for epoch 101: 0.15508951246738434
Validation accuracy for epoch 101: 0.9549000263214111
Epoch 102/200
Training loss for epoch 102: 0.14911207556724548
Training accuracy for epoch 102: 0.957770586013794
Validation loss for epoch 102: 0.1542753130197525
Validation accuracy for epoch 102: 0.9549000263214111
Epoch 103/200
Training loss for epoch 103: 0.14813625812530518
Training accuracy for epoch 103: 0.9580850601196289
Validation loss for epoch 103: 0.1534723937511444
Validation accuracy for epoch 103: 0.9549999833106995
Epoch 104/200
Training loss for epoch 104: 0.14717158675193787
Training accuracy for epoch 104: 0.9583830237388611
Validation loss for epoch 104: 0.15268047153949738
Validation accuracy for epoch 104: 0.955299973487854
Epoch 105/200
Training loss for epoch 105: 0.14621791243553162
Training accuracy for epoch 105: 0.9586312770843506
Validation loss for epoch 105: 0.15189936757087708
Validation accuracy for epoch 105: 0.9553999900817871
Epoch 106/200
Training loss for epoch 106: 0.14527501165866852
Training accuracy for epoch 106: 0.9587967991828918
Validation loss for epoch 106: 0.15112881362438202
Validation accuracy for epoch 106: 0.9555000066757202
Epoch 107/200
Training loss for epoch 107: 0.14434264600276947
Training accuracy for epoch 107: 0.9590451121330261

Validation loss for epoch 107: 0.1503685861825943
Validation accuracy for epoch 107: 0.9555000066757202
Epoch 108/200
Training loss for epoch 108: 0.14342063665390015
Training accuracy for epoch 108: 0.9592437148094177
Validation loss for epoch 108: 0.1496185064315796
Validation accuracy for epoch 108: 0.9553999900817871
Epoch 109/200
Training loss for epoch 109: 0.1425088495016098
Training accuracy for epoch 109: 0.9595251083374023
Validation loss for epoch 109: 0.14887835085391998
Validation accuracy for epoch 109: 0.9556999802589417
Epoch 110/200
Training loss for epoch 110: 0.14160704612731934
Training accuracy for epoch 110: 0.9599388837814331
Validation loss for epoch 110: 0.14814791083335876
Validation accuracy for epoch 110: 0.9557999968528748
Epoch 111/200
Training loss for epoch 111: 0.14071506261825562
Training accuracy for epoch 111: 0.9601374864578247
Validation loss for epoch 111: 0.14742697775363922
Validation accuracy for epoch 111: 0.9559999704360962
Epoch 112/200
Training loss for epoch 112: 0.1398327499628067
Training accuracy for epoch 112: 0.9605181813240051
Validation loss for epoch 112: 0.14671538770198822
Validation accuracy for epoch 112: 0.9564999938011169
Epoch 113/200
Training loss for epoch 113: 0.1389598846435547
Training accuracy for epoch 113: 0.9608327150344849
Validation loss for epoch 113: 0.14601291716098785
Validation accuracy for epoch 113: 0.9564999938011169
Epoch 114/200
Training loss for epoch 114: 0.13809634745121002
Training accuracy for epoch 114: 0.9611306190490723
Validation loss for epoch 114: 0.1453193873167038
Validation accuracy for epoch 114: 0.9567000269889832
Epoch 115/200
Training loss for epoch 115: 0.137241929769516
Training accuracy for epoch 115: 0.9614782333374023
Validation loss for epoch 115: 0.14463461935520172
Validation accuracy for epoch 115: 0.9567999839782715
Epoch 116/200
Training loss for epoch 116: 0.13639652729034424
Training accuracy for epoch 116: 0.9617761373519897
Validation loss for epoch 116: 0.1439584493637085
Validation accuracy for epoch 116: 0.95660001039505
Epoch 117/200

Training loss for epoch 117: 0.13555996119976044
Training accuracy for epoch 117: 0.9621734023094177
Validation loss for epoch 117: 0.1432906985282898
Validation accuracy for epoch 117: 0.9569000005722046
Epoch 118/200
Training loss for epoch 118: 0.13473205268383026
Training accuracy for epoch 118: 0.9624713063240051
Validation loss for epoch 118: 0.14263123273849487
Validation accuracy for epoch 118: 0.9570000171661377
Epoch 119/200
Training loss for epoch 119: 0.13391268253326416
Training accuracy for epoch 119: 0.9626368284225464
Validation loss for epoch 119: 0.14197982847690582
Validation accuracy for epoch 119: 0.9573000073432922
Epoch 120/200
Training loss for epoch 120: 0.1331017017364502
Training accuracy for epoch 120: 0.9628851413726807
Validation loss for epoch 120: 0.1413363516330719
Validation accuracy for epoch 120: 0.9574999809265137
Epoch 121/200
Training loss for epoch 121: 0.13229896128177643
Training accuracy for epoch 121: 0.9631168842315674
Validation loss for epoch 121: 0.14070065319538116
Validation accuracy for epoch 121: 0.9574999809265137
Epoch 122/200
Training loss for epoch 122: 0.13150432705879211
Training accuracy for epoch 122: 0.9633817076683044
Validation loss for epoch 122: 0.14007259905338287
Validation accuracy for epoch 122: 0.9574999809265137
Epoch 123/200
Training loss for epoch 123: 0.1307176649570465
Training accuracy for epoch 123: 0.963580310344696
Validation loss for epoch 123: 0.1394520252943039
Validation accuracy for epoch 123: 0.9577000141143799
Epoch 124/200
Training loss for epoch 124: 0.12993884086608887
Training accuracy for epoch 124: 0.9637954831123352
Validation loss for epoch 124: 0.1388387680053711
Validation accuracy for epoch 124: 0.9577999711036682
Epoch 125/200
Training loss for epoch 125: 0.12916772067546844
Training accuracy for epoch 125: 0.9640272259712219
Validation loss for epoch 125: 0.1382327377796173
Validation accuracy for epoch 125: 0.9578999876976013
Epoch 126/200
Training loss for epoch 126: 0.12840420007705688
Training accuracy for epoch 126: 0.9643086194992065
Validation loss for epoch 126: 0.1376337707042694

Validation accuracy for epoch 126: 0.9578999876976013
Epoch 127/200
Training loss for epoch 127: 0.12764813005924225
Training accuracy for epoch 127: 0.9644410014152527
Validation loss for epoch 127: 0.13704171776771545
Validation accuracy for epoch 127: 0.9578999876976013
Epoch 128/200
Training loss for epoch 128: 0.12689939141273499
Training accuracy for epoch 128: 0.9646396040916443
Validation loss for epoch 128: 0.1364564597606659
Validation accuracy for epoch 128: 0.9581000208854675
Epoch 129/200
Training loss for epoch 129: 0.12615787982940674
Training accuracy for epoch 129: 0.9648382663726807
Validation loss for epoch 129: 0.13587787747383118
Validation accuracy for epoch 129: 0.9585000276565552
Epoch 130/200
Training loss for epoch 130: 0.12542346119880676
Training accuracy for epoch 130: 0.9651030898094177
Validation loss for epoch 130: 0.13530588150024414
Validation accuracy for epoch 130: 0.9585999846458435
Epoch 131/200
Training loss for epoch 131: 0.1246960461139679
Training accuracy for epoch 131: 0.9653513431549072
Validation loss for epoch 131: 0.13474027812480927
Validation accuracy for epoch 131: 0.958899974822998
Epoch 132/200
Training loss for epoch 132: 0.1239754930138588
Training accuracy for epoch 132: 0.9656161665916443
Validation loss for epoch 132: 0.13418103754520416
Validation accuracy for epoch 132: 0.958899974822998
Epoch 133/200
Training loss for epoch 133: 0.12326173484325409
Training accuracy for epoch 133: 0.965847909450531
Validation loss for epoch 133: 0.13362795114517212
Validation accuracy for epoch 133: 0.9589999914169312
Epoch 134/200
Training loss for epoch 134: 0.12255464494228363
Training accuracy for epoch 134: 0.9660465717315674
Validation loss for epoch 134: 0.13308098912239075
Validation accuracy for epoch 134: 0.9589999914169312
Epoch 135/200
Training loss for epoch 135: 0.12185411155223846
Training accuracy for epoch 135: 0.966162383556366
Validation loss for epoch 135: 0.1325400173664093
Validation accuracy for epoch 135: 0.9592999815940857
Epoch 136/200
Training loss for epoch 136: 0.12116003781557083

Training accuracy for epoch 136: 0.9664106965065002
Validation loss for epoch 136: 0.13200490176677704
Validation accuracy for epoch 136: 0.9595999717712402
Epoch 137/200
Training loss for epoch 137: 0.12047234922647476
Training accuracy for epoch 137: 0.9666092991828918
Validation loss for epoch 137: 0.13147558271884918
Validation accuracy for epoch 137: 0.9598000049591064
Epoch 138/200
Training loss for epoch 138: 0.11979091167449951
Training accuracy for epoch 138: 0.9668079018592834
Validation loss for epoch 138: 0.13095195591449738
Validation accuracy for epoch 138: 0.9598000049591064
Epoch 139/200
Training loss for epoch 139: 0.11911564320325851
Training accuracy for epoch 139: 0.9670231342315674
Validation loss for epoch 139: 0.13043391704559326
Validation accuracy for epoch 139: 0.9599999785423279
Epoch 140/200
Training loss for epoch 140: 0.1184464693069458
Training accuracy for epoch 140: 0.9672383069992065
Validation loss for epoch 140: 0.1299213320016861
Validation accuracy for epoch 140: 0.9599999785423279
Epoch 141/200
Training loss for epoch 141: 0.11778329312801361
Training accuracy for epoch 141: 0.9674534797668457
Validation loss for epoch 141: 0.1294141560792923
Validation accuracy for epoch 141: 0.9603999853134155
Epoch 142/200
Training loss for epoch 142: 0.11712600290775299
Training accuracy for epoch 142: 0.9676024317741394
Validation loss for epoch 142: 0.1289122998714447
Validation accuracy for epoch 142: 0.9606000185012817
Epoch 143/200
Training loss for epoch 143: 0.11647452414035797
Training accuracy for epoch 143: 0.9678176045417786
Validation loss for epoch 143: 0.12841564416885376
Validation accuracy for epoch 143: 0.9606999754905701
Epoch 144/200
Training loss for epoch 144: 0.11582880467176437
Training accuracy for epoch 144: 0.9680162072181702
Validation loss for epoch 144: 0.1279241442680359
Validation accuracy for epoch 144: 0.9607999920845032
Epoch 145/200
Training loss for epoch 145: 0.11518870294094086
Training accuracy for epoch 145: 0.9681486487388611
Validation loss for epoch 145: 0.12743768095970154
Validation accuracy for epoch 145: 0.9609000086784363

Epoch 146/200
Training loss for epoch 146: 0.11455418169498444
Training accuracy for epoch 146: 0.9683306813240051
Validation loss for epoch 146: 0.12695619463920593
Validation accuracy for epoch 146: 0.9610999822616577
Epoch 147/200
Training loss for epoch 147: 0.11392515152692795
Training accuracy for epoch 147: 0.9684962034225464
Validation loss for epoch 147: 0.1264795958995819
Validation accuracy for epoch 147: 0.9611999988555908
Epoch 148/200
Training loss for epoch 148: 0.11330150067806244
Training accuracy for epoch 148: 0.9687445163726807
Validation loss for epoch 148: 0.1260078102350235
Validation accuracy for epoch 148: 0.9614999890327454
Epoch 149/200
Training loss for epoch 149: 0.11268322169780731
Training accuracy for epoch 149: 0.9688768982887268
Validation loss for epoch 149: 0.12554077804088593
Validation accuracy for epoch 149: 0.9617000222206116
Epoch 150/200
Training loss for epoch 150: 0.11207018047571182
Training accuracy for epoch 150: 0.969092071056366
Validation loss for epoch 150: 0.12507839500904083
Validation accuracy for epoch 150: 0.9617000222206116
Epoch 151/200
Training loss for epoch 151: 0.11146232485771179
Training accuracy for epoch 151: 0.9693403840065002
Validation loss for epoch 151: 0.12462061643600464
Validation accuracy for epoch 151: 0.9617000222206116
Epoch 152/200
Training loss for epoch 152: 0.11085961014032364
Training accuracy for epoch 152: 0.9694562554359436
Validation loss for epoch 152: 0.12416738271713257
Validation accuracy for epoch 152: 0.961899995803833
Epoch 153/200
Training loss for epoch 153: 0.11026190221309662
Training accuracy for epoch 153: 0.9696217775344849
Validation loss for epoch 153: 0.12371858209371567
Validation accuracy for epoch 153: 0.961899995803833
Epoch 154/200
Training loss for epoch 154: 0.10966917872428894
Training accuracy for epoch 154: 0.9697375893592834
Validation loss for epoch 154: 0.12327419221401215
Validation accuracy for epoch 154: 0.9621000289916992
Epoch 155/200
Training loss for epoch 155: 0.10908137261867523
Training accuracy for epoch 155: 0.9699196815490723

Validation loss for epoch 155: 0.12283414602279663
Validation accuracy for epoch 155: 0.9623000025749207
Epoch 156/200
Training loss for epoch 156: 0.10849840193986893
Training accuracy for epoch 156: 0.970068633556366
Validation loss for epoch 156: 0.12239833921194077
Validation accuracy for epoch 156: 0.9623000025749207
Epoch 157/200
Training loss for epoch 157: 0.10792020708322525
Training accuracy for epoch 157: 0.9702176451683044
Validation loss for epoch 157: 0.12196675688028336
Validation accuracy for epoch 157: 0.9624000191688538
Epoch 158/200
Training loss for epoch 158: 0.10734674334526062
Training accuracy for epoch 158: 0.9704493284225464
Validation loss for epoch 158: 0.12153930962085724
Validation accuracy for epoch 158: 0.9625999927520752
Epoch 159/200
Training loss for epoch 159: 0.10677789896726608
Training accuracy for epoch 159: 0.9706645011901855
Validation loss for epoch 159: 0.12111596763134003
Validation accuracy for epoch 159: 0.9628999829292297
Epoch 160/200
Training loss for epoch 160: 0.10621368885040283
Training accuracy for epoch 160: 0.9709128141403198
Validation loss for epoch 160: 0.12069665640592575
Validation accuracy for epoch 160: 0.963100016117096
Epoch 161/200
Training loss for epoch 161: 0.10565397888422012
Training accuracy for epoch 161: 0.9710948467254639
Validation loss for epoch 161: 0.12028130888938904
Validation accuracy for epoch 161: 0.9629999995231628
Epoch 162/200
Training loss for epoch 162: 0.10509877651929855
Training accuracy for epoch 162: 0.9712438583374023
Validation loss for epoch 162: 0.1198698952794075
Validation accuracy for epoch 162: 0.9628999829292297
Epoch 163/200
Training loss for epoch 163: 0.10454796999692917
Training accuracy for epoch 163: 0.9713762402534485
Validation loss for epoch 163: 0.11946235597133636
Validation accuracy for epoch 163: 0.9634000062942505
Epoch 164/200
Training loss for epoch 164: 0.10400152206420898
Training accuracy for epoch 164: 0.9715417623519897
Validation loss for epoch 164: 0.11905864626169205
Validation accuracy for epoch 164: 0.9634000062942505
Epoch 165/200

Training loss for epoch 165: 0.10345939546823502
Training accuracy for epoch 165: 0.9716742038726807
Validation loss for epoch 165: 0.11865869164466858
Validation accuracy for epoch 165: 0.9634000062942505
Epoch 166/200
Training loss for epoch 166: 0.10292152315378189
Training accuracy for epoch 166: 0.9718397259712219
Validation loss for epoch 166: 0.11826246231794357
Validation accuracy for epoch 166: 0.9634000062942505
Epoch 167/200
Training loss for epoch 167: 0.10238783806562424
Training accuracy for epoch 167: 0.972021758556366
Validation loss for epoch 167: 0.11786990612745285
Validation accuracy for epoch 167: 0.9635999798774719
Epoch 168/200
Training loss for epoch 168: 0.10185831040143967
Training accuracy for epoch 168: 0.972104549407959
Validation loss for epoch 168: 0.11748100817203522
Validation accuracy for epoch 168: 0.9635999798774719
Epoch 169/200
Training loss for epoch 169: 0.1013328805565834
Training accuracy for epoch 169: 0.9721872806549072
Validation loss for epoch 169: 0.11709565669298172
Validation accuracy for epoch 169: 0.9638000130653381
Epoch 170/200
Training loss for epoch 170: 0.10081151127815247
Training accuracy for epoch 170: 0.9723197221755981
Validation loss for epoch 170: 0.11671384423971176
Validation accuracy for epoch 170: 0.9639000296592712
Epoch 171/200
Training loss for epoch 171: 0.10029412060976028
Training accuracy for epoch 171: 0.9725679755210876
Validation loss for epoch 171: 0.11633552610874176
Validation accuracy for epoch 171: 0.9639000296592712
Epoch 172/200
Training loss for epoch 172: 0.09978070855140686
Training accuracy for epoch 172: 0.972683846950531
Validation loss for epoch 172: 0.11596065759658813
Validation accuracy for epoch 172: 0.9639000296592712
Epoch 173/200
Training loss for epoch 173: 0.09927117824554443
Training accuracy for epoch 173: 0.9728327989578247
Validation loss for epoch 173: 0.1155892089009285
Validation accuracy for epoch 173: 0.9642000198364258
Epoch 174/200
Training loss for epoch 174: 0.0987655371427536
Training accuracy for epoch 174: 0.9729321599006653
Validation loss for epoch 174: 0.11522110551595688

Validation accuracy for epoch 174: 0.9643999934196472
Epoch 175/200
Training loss for epoch 175: 0.0982636883854866
Training accuracy for epoch 175: 0.9729486703872681
Validation loss for epoch 175: 0.11485634744167328
Validation accuracy for epoch 175: 0.9646999835968018
Epoch 176/200
Training loss for epoch 176: 0.09776560962200165
Training accuracy for epoch 176: 0.9731307625770569
Validation loss for epoch 176: 0.11449487507343292
Validation accuracy for epoch 176: 0.964900016784668
Epoch 177/200
Training loss for epoch 177: 0.09727126359939575
Training accuracy for epoch 177: 0.9732466340065002
Validation loss for epoch 177: 0.11413665860891342
Validation accuracy for epoch 177: 0.964900016784668
Epoch 178/200
Training loss for epoch 178: 0.09678059071302414
Training accuracy for epoch 178: 0.9733790159225464
Validation loss for epoch 178: 0.1137816533446312
Validation accuracy for epoch 178: 0.9650999903678894
Epoch 179/200
Training loss for epoch 179: 0.09629358351230621
Training accuracy for epoch 179: 0.9735776782035828
Validation loss for epoch 179: 0.11342983692884445
Validation accuracy for epoch 179: 0.9650999903678894
Epoch 180/200
Training loss for epoch 180: 0.0958101823925972
Training accuracy for epoch 180: 0.9736769795417786
Validation loss for epoch 180: 0.11308112740516663
Validation accuracy for epoch 180: 0.9650999903678894
Epoch 181/200
Training loss for epoch 181: 0.09533033519983292
Training accuracy for epoch 181: 0.9738259315490723
Validation loss for epoch 181: 0.11273554712533951
Validation accuracy for epoch 181: 0.9652000069618225
Epoch 182/200
Training loss for epoch 182: 0.09485399723052979
Training accuracy for epoch 182: 0.974057674407959
Validation loss for epoch 182: 0.11239302903413773
Validation accuracy for epoch 182: 0.9653000235557556
Epoch 183/200
Training loss for epoch 183: 0.09438114613294601
Training accuracy for epoch 183: 0.9742066264152527
Validation loss for epoch 183: 0.11205354332923889
Validation accuracy for epoch 183: 0.9653000235557556
Epoch 184/200
Training loss for epoch 184: 0.09391175210475922

Training accuracy for epoch 184: 0.9744052290916443
Validation loss for epoch 184: 0.11171707510948181
Validation accuracy for epoch 184: 0.9653000235557556
Epoch 185/200
Training loss for epoch 185: 0.09344575554132462
Training accuracy for epoch 185: 0.9746038913726807
Validation loss for epoch 185: 0.11138357222080231
Validation accuracy for epoch 185: 0.965399980545044
Epoch 186/200
Training loss for epoch 186: 0.09298313409090042
Training accuracy for epoch 186: 0.974719762802124
Validation loss for epoch 186: 0.11105300486087799
Validation accuracy for epoch 186: 0.9656999707221985
Epoch 187/200
Training loss for epoch 187: 0.09252384305000305
Training accuracy for epoch 187: 0.9748521447181702
Validation loss for epoch 187: 0.11072536557912827
Validation accuracy for epoch 187: 0.9656999707221985
Epoch 188/200
Training loss for epoch 188: 0.09206784516572952
Training accuracy for epoch 188: 0.9749680161476135
Validation loss for epoch 188: 0.11040058732032776
Validation accuracy for epoch 188: 0.9657999873161316
Epoch 189/200
Training loss for epoch 189: 0.09161512553691864
Training accuracy for epoch 189: 0.9750673174858093
Validation loss for epoch 189: 0.11007864773273468
Validation accuracy for epoch 189: 0.9659000039100647
Epoch 190/200
Training loss for epoch 190: 0.09116563200950623
Training accuracy for epoch 190: 0.9751666188240051
Validation loss for epoch 190: 0.10975954681634903
Validation accuracy for epoch 190: 0.9660000205039978
Epoch 191/200
Training loss for epoch 191: 0.0907193273305893
Training accuracy for epoch 191: 0.9752659797668457
Validation loss for epoch 191: 0.10944323986768723
Validation accuracy for epoch 191: 0.9660999774932861
Epoch 192/200
Training loss for epoch 192: 0.09027618914842606
Training accuracy for epoch 192: 0.9754645824432373
Validation loss for epoch 192: 0.1091296598315239
Validation accuracy for epoch 192: 0.9660999774932861
Epoch 193/200
Training loss for epoch 193: 0.08983617275953293
Training accuracy for epoch 193: 0.9755804538726807
Validation loss for epoch 193: 0.10881884396076202
Validation accuracy for epoch 193: 0.9661999940872192

```

Epoch 194/200
Training loss for epoch 194: 0.08939925581216812
Training accuracy for epoch 194: 0.9757624864578247
Validation loss for epoch 194: 0.10851071774959564
Validation accuracy for epoch 194: 0.9661999940872192
Epoch 195/200
Training loss for epoch 195: 0.08896540850400925
Training accuracy for epoch 195: 0.975928008556366
Validation loss for epoch 195: 0.10820527374744415
Validation accuracy for epoch 195: 0.9664000272750854
Epoch 196/200
Training loss for epoch 196: 0.08853457868099213
Training accuracy for epoch 196: 0.9760935306549072
Validation loss for epoch 196: 0.10790248960256577
Validation accuracy for epoch 196: 0.9664000272750854
Epoch 197/200
Training loss for epoch 197: 0.08810675889253616
Training accuracy for epoch 197: 0.9762094020843506
Validation loss for epoch 197: 0.1076023206114769
Validation accuracy for epoch 197: 0.9666000008583069
Epoch 198/200
Training loss for epoch 198: 0.08768191933631897
Training accuracy for epoch 198: 0.9763087034225464
Validation loss for epoch 198: 0.10730475932359695
Validation accuracy for epoch 198: 0.9668999910354614
Epoch 199/200
Training loss for epoch 199: 0.08726000040769577
Training accuracy for epoch 199: 0.9764411449432373
Validation loss for epoch 199: 0.10700976848602295
Validation accuracy for epoch 199: 0.9668999910354614
Epoch 200/200
Training loss for epoch 200: 0.08684101700782776
Training accuracy for epoch 200: 0.9765238761901855
Validation loss for epoch 200: 0.1067173108458519
Validation accuracy for epoch 200: 0.9670000076293945

```

5 Gradient Accumulation Model

```

[ ]: # Initialize lists to store metrics
train_losses = []
val_losses = []
train_accuracies = []
val_accuracies = []

# Number of epochs to train for
num_epochs = 200

```

```

# Initialize accumulators
accumulated_grads = [tf.zeros_like(w) for w in a_model.trainable_weights]

for epoch in range(num_epochs):
    print(f"Epoch {epoch+1}/{num_epochs}")

    # Initialize lists to store metrics for each batch
    batch_losses = []
    batch_accuracies = []

    # Iterate over the batches of the dataset
    for step, (X_batch, y_batch) in enumerate(data_mnist):
        with tf.GradientTape() as tape:
            # Forward pass
            y_pred = a_model(X_batch, training=True)
            # Compute the loss value
            loss_value = loss_fn(y_batch, y_pred)

            # Compute gradients - backward pass
            grads = tape.gradient(loss_value, a_model.trainable_weights)

            # If it's the final batch, adjust the gradients
            if (step + 1) == len(data_mnist):
                final_batch_size = len(X_batch)
                grads = [g * final_batch_size / batch_size for g in grads]

            # Accumulate gradients
            accumulated_grads = [acc_g + g for acc_g, g in zip(accumulated_grads,
↪grads)]

            # Store metrics for this batch
            batch_losses.append(loss_value.numpy())
            batch_accuracies.append(tf.reduce_mean(tf.cast(tf.equal(tf.
↪cast(y_batch, tf.int64), tf.argmax(y_pred, axis=1)), tf.float32)))

        # Normalize the accumulated gradients
        accumulated_grads = [acc_g / len(data_mnist) for acc_g in accumulated_grads]

    # Update weights after all batches are processed
    optimizer.apply_gradients(zip(accumulated_grads, a_model.trainable_weights))

    # Compute average metrics for this epoch
    avg_train_loss = np.mean(batch_losses)
    avg_train_accuracy = np.mean(batch_accuracies)

    # Print the loss and accuracy for each epoch
    print(f"Training loss for epoch {epoch+1}: {avg_train_loss}")

```

```

print(f"Training accuracy for epoch {epoch+1}: {avg_train_accuracy}")

# Append average metrics for this epoch
train_losses.append(avg_train_loss)
train_accuracies.append(avg_train_accuracy)

# Compute validation metrics for this epoch
y_val_pred = a_model(X_test_mnist, training=False)
val_loss = np.mean(loss_fn(y_test_mnist, y_val_pred).numpy())
val_accuracy = np.mean(tf.cast(tf.equal(y_test_mnist, tf.argmax(y_val_pred, ↵
↵axis=1))), tf.float32).numpy())

val_losses.append(val_loss)
val_accuracies.append(val_accuracy)

# Reset the accumulators for the next epoch
accumulated_grads = [tf.zeros_like(w) for w in a_model.trainable_weights]

```

```

Epoch 1/200
Training loss for epoch 1: 2.3934335708618164
Training accuracy for epoch 1: 0.11590748280286789
Epoch 2/200
Training loss for epoch 2: 2.368044853210449
Training accuracy for epoch 2: 0.12971177697181702
Epoch 3/200
Training loss for epoch 3: 2.343747615814209
Training accuracy for epoch 3: 0.14344985783100128
Epoch 4/200
Training loss for epoch 4: 2.320418357849121
Training accuracy for epoch 4: 0.15730932354927063
Epoch 5/200
Training loss for epoch 5: 2.2979559898376465
Training accuracy for epoch 5: 0.17137844860553741
Epoch 6/200
Training loss for epoch 6: 2.2762787342071533
Training accuracy for epoch 6: 0.185618594288826
Epoch 7/200
Training loss for epoch 7: 2.2553164958953857
Training accuracy for epoch 7: 0.20054841041564941
Epoch 8/200
Training loss for epoch 8: 2.2350122928619385
Training accuracy for epoch 8: 0.21458442509174347
Epoch 9/200
Training loss for epoch 9: 2.2153170108795166
Training accuracy for epoch 9: 0.22819562256336212
Epoch 10/200
Training loss for epoch 10: 2.196187734603882
Training accuracy for epoch 10: 0.2415364682674408

```

Epoch 11/200
Training loss for epoch 11: 2.1775875091552734
Training accuracy for epoch 11: 0.25587591528892517
Epoch 12/200
Training loss for epoch 12: 2.159482717514038
Training accuracy for epoch 12: 0.26908984780311584
Epoch 13/200
Training loss for epoch 13: 2.14184308052063
Training accuracy for epoch 13: 0.28173553943634033
Epoch 14/200
Training loss for epoch 14: 2.124640703201294
Training accuracy for epoch 14: 0.29368600249290466
Epoch 15/200
Training loss for epoch 15: 2.107851266860962
Training accuracy for epoch 15: 0.3048364818096161
Epoch 16/200
Training loss for epoch 16: 2.0914504528045654
Training accuracy for epoch 16: 0.3161083161830902
Epoch 17/200
Training loss for epoch 17: 2.0754170417785645
Training accuracy for epoch 17: 0.3270767033100128
Epoch 18/200
Training loss for epoch 18: 2.0597317218780518
Training accuracy for epoch 18: 0.3384478688240051
Epoch 19/200
Training loss for epoch 19: 2.0443758964538574
Training accuracy for epoch 19: 0.3492617905139923
Epoch 20/200
Training loss for epoch 20: 2.029332399368286
Training accuracy for epoch 20: 0.35933637619018555
Epoch 21/200
Training loss for epoch 21: 2.0145862102508545
Training accuracy for epoch 21: 0.3694220185279846
Epoch 22/200
Training loss for epoch 22: 2.0001230239868164
Training accuracy for epoch 22: 0.37939178943634033
Epoch 23/200
Training loss for epoch 23: 1.9859297275543213
Training accuracy for epoch 23: 0.3890470266342163
Epoch 24/200
Training loss for epoch 24: 1.9719946384429932
Training accuracy for epoch 24: 0.39854782819747925
Epoch 25/200
Training loss for epoch 25: 1.9583070278167725
Training accuracy for epoch 25: 0.40781691670417786
Epoch 26/200
Training loss for epoch 26: 1.9448565244674683
Training accuracy for epoch 26: 0.4159384071826935

Epoch 27/200
Training loss for epoch 27: 1.9316350221633911
Training accuracy for epoch 27: 0.4252074658870697
Epoch 28/200
Training loss for epoch 28: 1.918632984161377
Training accuracy for epoch 28: 0.4337702989578247
Epoch 29/200
Training loss for epoch 29: 1.9058436155319214
Training accuracy for epoch 29: 0.4423883259296417
Epoch 30/200
Training loss for epoch 30: 1.89326012134552
Training accuracy for epoch 30: 0.450564980506897
Epoch 31/200
Training loss for epoch 31: 1.880875825881958
Training accuracy for epoch 31: 0.4584602415561676
Epoch 32/200
Training loss for epoch 32: 1.8686846494674683
Training accuracy for epoch 32: 0.4666423797607422
Epoch 33/200
Training loss for epoch 33: 1.8566814661026
Training accuracy for epoch 33: 0.4741404056549072
Epoch 34/200
Training loss for epoch 34: 1.844861388206482
Training accuracy for epoch 34: 0.481274276971817
Epoch 35/200
Training loss for epoch 35: 1.833219051361084
Training accuracy for epoch 35: 0.4885791838169098
Epoch 36/200
Training loss for epoch 36: 1.8217509984970093
Training accuracy for epoch 36: 0.4951392710208893
Epoch 37/200
Training loss for epoch 37: 1.8104525804519653
Training accuracy for epoch 37: 0.5016165375709534
Epoch 38/200
Training loss for epoch 38: 1.7993203401565552
Training accuracy for epoch 38: 0.5072993636131287
Epoch 39/200
Training loss for epoch 39: 1.7883504629135132
Training accuracy for epoch 39: 0.5127118825912476
Epoch 40/200
Training loss for epoch 40: 1.7775393724441528
Training accuracy for epoch 40: 0.5179974436759949
Epoch 41/200
Training loss for epoch 41: 1.7668839693069458
Training accuracy for epoch 41: 0.5239726901054382
Epoch 42/200
Training loss for epoch 42: 1.7563813924789429
Training accuracy for epoch 42: 0.5293024182319641

Epoch 43/200
Training loss for epoch 43: 1.7460283041000366
Training accuracy for epoch 43: 0.534234881401062
Epoch 44/200
Training loss for epoch 44: 1.7358224391937256
Training accuracy for epoch 44: 0.539432168006897
Epoch 45/200
Training loss for epoch 45: 1.7257602214813232
Training accuracy for epoch 45: 0.5442377328872681
Epoch 46/200
Training loss for epoch 46: 1.7158398628234863
Training accuracy for epoch 46: 0.5491922497749329
Epoch 47/200
Training loss for epoch 47: 1.706058144569397
Training accuracy for epoch 47: 0.5538157224655151
Epoch 48/200
Training loss for epoch 48: 1.696413516998291
Training accuracy for epoch 48: 0.5585827231407166
Epoch 49/200
Training loss for epoch 49: 1.6869022846221924
Training accuracy for epoch 49: 0.5630020499229431
Epoch 50/200
Training loss for epoch 50: 1.677523136138916
Training accuracy for epoch 50: 0.567040741443634
Epoch 51/200
Training loss for epoch 51: 1.6682735681533813
Training accuracy for epoch 51: 0.5707152485847473
Epoch 52/200
Training loss for epoch 52: 1.659151315689087
Training accuracy for epoch 52: 0.5748697519302368
Epoch 53/200
Training loss for epoch 53: 1.6501542329788208
Training accuracy for epoch 53: 0.5790077447891235
Epoch 54/200
Training loss for epoch 54: 1.6412800550460815
Training accuracy for epoch 54: 0.5825663805007935
Epoch 55/200
Training loss for epoch 55: 1.6325267553329468
Training accuracy for epoch 55: 0.5861912965774536
Epoch 56/200
Training loss for epoch 56: 1.6238924264907837
Training accuracy for epoch 56: 0.5896174907684326
Epoch 57/200
Training loss for epoch 57: 1.6153748035430908
Training accuracy for epoch 57: 0.5929940938949585
Epoch 58/200
Training loss for epoch 58: 1.6069722175598145
Training accuracy for epoch 58: 0.5962382555007935

Epoch 59/200
Training loss for epoch 59: 1.5986822843551636
Training accuracy for epoch 59: 0.599454939365387
Epoch 60/200
Training loss for epoch 60: 1.5905038118362427
Training accuracy for epoch 60: 0.602384626865387
Epoch 61/200
Training loss for epoch 61: 1.5824344158172607
Training accuracy for epoch 61: 0.6057446599006653
Epoch 62/200
Training loss for epoch 62: 1.574472188949585
Training accuracy for epoch 62: 0.6084426045417786
Epoch 63/200
Training loss for epoch 63: 1.5666154623031616
Training accuracy for epoch 63: 0.6113446950912476
Epoch 64/200
Training loss for epoch 64: 1.5588622093200684
Training accuracy for epoch 64: 0.6139598488807678
Epoch 65/200
Training loss for epoch 65: 1.5512109994888306
Training accuracy for epoch 65: 0.6169668436050415
Epoch 66/200
Training loss for epoch 66: 1.5436598062515259
Training accuracy for epoch 66: 0.619168221950531
Epoch 67/200
Training loss for epoch 67: 1.5362074375152588
Training accuracy for epoch 67: 0.6217172145843506
Epoch 68/200
Training loss for epoch 68: 1.5288511514663696
Training accuracy for epoch 68: 0.6245972514152527
Epoch 69/200
Training loss for epoch 69: 1.5215903520584106
Training accuracy for epoch 69: 0.6272510886192322
Epoch 70/200
Training loss for epoch 70: 1.51442289352417
Training accuracy for epoch 70: 0.6294193863868713
Epoch 71/200
Training loss for epoch 71: 1.5073471069335938
Training accuracy for epoch 71: 0.6316373348236084
Epoch 72/200
Training loss for epoch 72: 1.500361442565918
Training accuracy for epoch 72: 0.6339876651763916
Epoch 73/200
Training loss for epoch 73: 1.493464469909668
Training accuracy for epoch 73: 0.6360621452331543
Epoch 74/200
Training loss for epoch 74: 1.4866544008255005
Training accuracy for epoch 74: 0.638517439365387

Epoch 75/200
Training loss for epoch 75: 1.4799299240112305
Training accuracy for epoch 75: 0.6404098272323608
Epoch 76/200
Training loss for epoch 76: 1.4732896089553833
Training accuracy for epoch 76: 0.6426112055778503
Epoch 77/200
Training loss for epoch 77: 1.4667317867279053
Training accuracy for epoch 77: 0.645165741443634
Epoch 78/200
Training loss for epoch 78: 1.4602547883987427
Training accuracy for epoch 78: 0.6472402811050415
Epoch 79/200
Training loss for epoch 79: 1.4538575410842896
Training accuracy for epoch 79: 0.6492927074432373
Epoch 80/200
Training loss for epoch 80: 1.44753897190094
Training accuracy for epoch 80: 0.6511630415916443
Epoch 81/200
Training loss for epoch 81: 1.4412968158721924
Training accuracy for epoch 81: 0.6530500054359436
Epoch 82/200
Training loss for epoch 82: 1.4351304769515991
Training accuracy for epoch 82: 0.6548541188240051
Epoch 83/200
Training loss for epoch 83: 1.4290380477905273
Training accuracy for epoch 83: 0.6565093398094177
Epoch 84/200
Training loss for epoch 84: 1.4230185747146606
Training accuracy for epoch 84: 0.6581976413726807
Epoch 85/200
Training loss for epoch 85: 1.4170706272125244
Training accuracy for epoch 85: 0.6596211194992065
Epoch 86/200
Training loss for epoch 86: 1.4111928939819336
Training accuracy for epoch 86: 0.6612597107887268
Epoch 87/200
Training loss for epoch 87: 1.4053843021392822
Training accuracy for epoch 87: 0.6629480123519897
Epoch 88/200
Training loss for epoch 88: 1.3996434211730957
Training accuracy for epoch 88: 0.6649397611618042
Epoch 89/200
Training loss for epoch 89: 1.3939695358276367
Training accuracy for epoch 89: 0.6666114926338196
Epoch 90/200
Training loss for epoch 90: 1.3883609771728516
Training accuracy for epoch 90: 0.6682170033454895

Epoch 91/200
Training loss for epoch 91: 1.3828169107437134
Training accuracy for epoch 91: 0.6697232723236084
Epoch 92/200
Training loss for epoch 92: 1.377335786819458
Training accuracy for epoch 92: 0.6713950037956238
Epoch 93/200
Training loss for epoch 93: 1.371916651725769
Training accuracy for epoch 93: 0.6731991767883301
Epoch 94/200
Training loss for epoch 94: 1.3665587902069092
Training accuracy for epoch 94: 0.6745564341545105
Epoch 95/200
Training loss for epoch 95: 1.3612607717514038
Training accuracy for epoch 95: 0.6759522557258606
Epoch 96/200
Training loss for epoch 96: 1.3560216426849365
Training accuracy for epoch 96: 0.6776571273803711
Epoch 97/200
Training loss for epoch 97: 1.350840449333191
Training accuracy for epoch 97: 0.6792847514152527
Epoch 98/200
Training loss for epoch 98: 1.345715880393982
Training accuracy for epoch 98: 0.6809564828872681
Epoch 99/200
Training loss for epoch 99: 1.3406476974487305
Training accuracy for epoch 99: 0.682727575302124
Epoch 100/200
Training loss for epoch 100: 1.3356341123580933
Training accuracy for epoch 100: 0.6843496561050415
Epoch 101/200
Training loss for epoch 101: 1.330674648284912
Training accuracy for epoch 101: 0.6858724355697632
Epoch 102/200
Training loss for epoch 102: 1.3257678747177124
Training accuracy for epoch 102: 0.6873786449432373
Epoch 103/200
Training loss for epoch 103: 1.320913553237915
Training accuracy for epoch 103: 0.6885703802108765
Epoch 104/200
Training loss for epoch 104: 1.3161104917526245
Training accuracy for epoch 104: 0.6898945569992065
Epoch 105/200
Training loss for epoch 105: 1.311358094215393
Training accuracy for epoch 105: 0.6913179755210876
Epoch 106/200
Training loss for epoch 106: 1.306654691696167
Training accuracy for epoch 106: 0.6925759315490723

Epoch 107/200
Training loss for epoch 107: 1.3020004034042358
Training accuracy for epoch 107: 0.6939994096755981
Epoch 108/200
Training loss for epoch 108: 1.2973936796188354
Training accuracy for epoch 108: 0.6952573657035828
Epoch 109/200
Training loss for epoch 109: 1.2928342819213867
Training accuracy for epoch 109: 0.6964711546897888
Epoch 110/200
Training loss for epoch 110: 1.2883211374282837
Training accuracy for epoch 110: 0.6977952718734741
Epoch 111/200
Training loss for epoch 111: 1.283853530883789
Training accuracy for epoch 111: 0.6991028785705566
Epoch 112/200
Training loss for epoch 112: 1.2794307470321655
Training accuracy for epoch 112: 0.7004435658454895
Epoch 113/200
Training loss for epoch 113: 1.2750520706176758
Training accuracy for epoch 113: 0.7017015218734741
Epoch 114/200
Training loss for epoch 114: 1.2707165479660034
Training accuracy for epoch 114: 0.702727735042572
Epoch 115/200
Training loss for epoch 115: 1.2664237022399902
Training accuracy for epoch 115: 0.7039029598236084
Epoch 116/200
Training loss for epoch 116: 1.2621731758117676
Training accuracy for epoch 116: 0.7052767276763916
Epoch 117/200
Training loss for epoch 117: 1.2579634189605713
Training accuracy for epoch 117: 0.7064353823661804
Epoch 118/200
Training loss for epoch 118: 1.2537944316864014
Training accuracy for epoch 118: 0.7076271176338196
Epoch 119/200
Training loss for epoch 119: 1.2496654987335205
Training accuracy for epoch 119: 0.7090174555778503
Epoch 120/200
Training loss for epoch 120: 1.2455755472183228
Training accuracy for epoch 120: 0.7101761102676392
Epoch 121/200
Training loss for epoch 121: 1.2415246963500977
Training accuracy for epoch 121: 0.7114837169647217
Epoch 122/200
Training loss for epoch 122: 1.2375116348266602
Training accuracy for epoch 122: 0.7127416729927063

Epoch 123/200
Training loss for epoch 123: 1.2335361242294312
Training accuracy for epoch 123: 0.7135361433029175
Epoch 124/200
Training loss for epoch 124: 1.229597568511963
Training accuracy for epoch 124: 0.7147278785705566
Epoch 125/200
Training loss for epoch 125: 1.2256953716278076
Training accuracy for epoch 125: 0.7158533930778503
Epoch 126/200
Training loss for epoch 126: 1.221828818321228
Training accuracy for epoch 126: 0.7169954776763916
Epoch 127/200
Training loss for epoch 127: 1.2179975509643555
Training accuracy for epoch 127: 0.7180714011192322
Epoch 128/200
Training loss for epoch 128: 1.2142008543014526
Training accuracy for epoch 128: 0.7192796468734741
Epoch 129/200
Training loss for epoch 129: 1.2104384899139404
Training accuracy for epoch 129: 0.7205045223236084
Epoch 130/200
Training loss for epoch 130: 1.2067095041275024
Training accuracy for epoch 130: 0.7216245532035828
Epoch 131/200
Training loss for epoch 131: 1.2030138969421387
Training accuracy for epoch 131: 0.7226065993309021
Epoch 132/200
Training loss for epoch 132: 1.1993508338928223
Training accuracy for epoch 132: 0.7235335111618042
Epoch 133/200
Training loss for epoch 133: 1.1957199573516846
Training accuracy for epoch 133: 0.7244438529014587
Epoch 134/200
Training loss for epoch 134: 1.1921207904815674
Training accuracy for epoch 134: 0.7256355881690979
Epoch 135/200
Training loss for epoch 135: 1.1885526180267334
Training accuracy for epoch 135: 0.7265294194221497
Epoch 136/200
Training loss for epoch 136: 1.1850149631500244
Training accuracy for epoch 136: 0.7275887131690979
Epoch 137/200
Training loss for epoch 137: 1.1815080642700195
Training accuracy for epoch 137: 0.7286645770072937
Epoch 138/200
Training loss for epoch 138: 1.1780307292938232
Training accuracy for epoch 138: 0.7296135425567627

Epoch 139/200
Training loss for epoch 139: 1.174583077430725
Training accuracy for epoch 139: 0.7304245829582214
Epoch 140/200
Training loss for epoch 140: 1.1711642742156982
Training accuracy for epoch 140: 0.7313184142112732
Epoch 141/200
Training loss for epoch 141: 1.1677738428115845
Training accuracy for epoch 141: 0.732079803943634
Epoch 142/200
Training loss for epoch 142: 1.1644116640090942
Training accuracy for epoch 142: 0.733056366443634
Epoch 143/200
Training loss for epoch 143: 1.1610773801803589
Training accuracy for epoch 143: 0.7341653108596802
Epoch 144/200
Training loss for epoch 144: 1.1577701568603516
Training accuracy for epoch 144: 0.7351584434509277
Epoch 145/200
Training loss for epoch 145: 1.154490351676941
Training accuracy for epoch 145: 0.7361018657684326
Epoch 146/200
Training loss for epoch 146: 1.1512370109558105
Training accuracy for epoch 146: 0.7368963956832886
Epoch 147/200
Training loss for epoch 147: 1.1480097770690918
Training accuracy for epoch 147: 0.7378895282745361
Epoch 148/200
Training loss for epoch 148: 1.1448086500167847
Training accuracy for epoch 148: 0.7387005686759949
Epoch 149/200
Training loss for epoch 149: 1.1416326761245728
Training accuracy for epoch 149: 0.7395116090774536
Epoch 150/200
Training loss for epoch 150: 1.1384824514389038
Training accuracy for epoch 150: 0.7404053807258606
Epoch 151/200
Training loss for epoch 151: 1.1353565454483032
Training accuracy for epoch 151: 0.7412495613098145
Epoch 152/200
Training loss for epoch 152: 1.1322555541992188
Training accuracy for epoch 152: 0.7421102523803711
Epoch 153/200
Training loss for epoch 153: 1.129178524017334
Training accuracy for epoch 153: 0.7429709434509277
Epoch 154/200
Training loss for epoch 154: 1.126125454902649
Training accuracy for epoch 154: 0.744030237197876

Epoch 155/200
Training loss for epoch 155: 1.1230958700180054
Training accuracy for epoch 155: 0.7448247671127319
Epoch 156/200
Training loss for epoch 156: 1.1200894117355347
Training accuracy for epoch 156: 0.7455530166625977
Epoch 157/200
Training loss for epoch 157: 1.1171057224273682
Training accuracy for epoch 157: 0.746330976486206
Epoch 158/200
Training loss for epoch 158: 1.1141449213027954
Training accuracy for epoch 158: 0.7470757961273193
Epoch 159/200
Training loss for epoch 159: 1.1112064123153687
Training accuracy for epoch 159: 0.7478537559509277
Epoch 160/200
Training loss for epoch 160: 1.1082898378372192
Training accuracy for epoch 160: 0.748681366443634
Epoch 161/200
Training loss for epoch 161: 1.1053950786590576
Training accuracy for epoch 161: 0.7494924068450928
Epoch 162/200
Training loss for epoch 162: 1.1025217771530151
Training accuracy for epoch 162: 0.7501710057258606
Epoch 163/200
Training loss for epoch 163: 1.0996695756912231
Training accuracy for epoch 163: 0.7508165240287781
Epoch 164/200
Training loss for epoch 164: 1.0968382358551025
Training accuracy for epoch 164: 0.7514123916625977
Epoch 165/200
Training loss for epoch 165: 1.0940277576446533
Training accuracy for epoch 165: 0.7523062229156494
Epoch 166/200
Training loss for epoch 166: 1.0912377834320068
Training accuracy for epoch 166: 0.7531006932258606
Epoch 167/200
Training loss for epoch 167: 1.0884677171707153
Training accuracy for epoch 167: 0.7538290023803711
Epoch 168/200
Training loss for epoch 168: 1.0857176780700684
Training accuracy for epoch 168: 0.754540741443634
Epoch 169/200
Training loss for epoch 169: 1.0829873085021973
Training accuracy for epoch 169: 0.7553021311759949
Epoch 170/200
Training loss for epoch 170: 1.0802762508392334
Training accuracy for epoch 170: 0.7561628222465515

Epoch 171/200
Training loss for epoch 171: 1.0775843858718872
Training accuracy for epoch 171: 0.756808340549469
Epoch 172/200
Training loss for epoch 172: 1.07491135597229
Training accuracy for epoch 172: 0.7575697302818298
Epoch 173/200
Training loss for epoch 173: 1.072257399559021
Training accuracy for epoch 173: 0.758099377155304
Epoch 174/200
Training loss for epoch 174: 1.0696215629577637
Training accuracy for epoch 174: 0.7587780356407166
Epoch 175/200
Training loss for epoch 175: 1.0670043230056763
Training accuracy for epoch 175: 0.7596221566200256
Epoch 176/200
Training loss for epoch 176: 1.064404845237732
Training accuracy for epoch 176: 0.7603504657745361
Epoch 177/200
Training loss for epoch 177: 1.0618232488632202
Training accuracy for epoch 177: 0.7608469724655151
Epoch 178/200
Training loss for epoch 178: 1.0592594146728516
Training accuracy for epoch 178: 0.7615587115287781
Epoch 179/200
Training loss for epoch 179: 1.0567128658294678
Training accuracy for epoch 179: 0.7622207999229431
Epoch 180/200
Training loss for epoch 180: 1.0541836023330688
Training accuracy for epoch 180: 0.7631036043167114
Epoch 181/200
Training loss for epoch 181: 1.0516711473464966
Training accuracy for epoch 181: 0.7638815641403198
Epoch 182/200
Training loss for epoch 182: 1.0491758584976196
Training accuracy for epoch 182: 0.7645933032035828
Epoch 183/200
Training loss for epoch 183: 1.0466970205307007
Training accuracy for epoch 183: 0.7651450037956238
Epoch 184/200
Training loss for epoch 184: 1.0442345142364502
Training accuracy for epoch 184: 0.765939474105835
Epoch 185/200
Training loss for epoch 185: 1.0417882204055786
Training accuracy for epoch 185: 0.7665684819221497
Epoch 186/200
Training loss for epoch 186: 1.0393582582473755
Training accuracy for epoch 186: 0.7671974301338196

Epoch 187/200
Training loss for epoch 187: 1.036944031715393
Training accuracy for epoch 187: 0.7679588198661804
Epoch 188/200
Training loss for epoch 188: 1.0345455408096313
Training accuracy for epoch 188: 0.7687588334083557
Epoch 189/200
Training loss for epoch 189: 1.0321624279022217
Training accuracy for epoch 189: 0.7694043517112732
Epoch 190/200
Training loss for epoch 190: 1.0297948122024536
Training accuracy for epoch 190: 0.7698843479156494
Epoch 191/200
Training loss for epoch 191: 1.0274423360824585
Training accuracy for epoch 191: 0.7705464363098145
Epoch 192/200
Training loss for epoch 192: 1.0251049995422363
Training accuracy for epoch 192: 0.7711091637611389
Epoch 193/200
Training loss for epoch 193: 1.0227824449539185
Training accuracy for epoch 193: 0.7717546820640564
Epoch 194/200
Training loss for epoch 194: 1.0204744338989258
Training accuracy for epoch 194: 0.7723174691200256
Epoch 195/200
Training loss for epoch 195: 1.0181810855865479
Training accuracy for epoch 195: 0.7729299068450928
Epoch 196/200
Training loss for epoch 196: 1.0159021615982056
Training accuracy for epoch 196: 0.7736968398094177
Epoch 197/200
Training loss for epoch 197: 1.0136373043060303
Training accuracy for epoch 197: 0.7744140625
Epoch 198/200
Training loss for epoch 198: 1.0113866329193115
Training accuracy for epoch 198: 0.7750265002250671
Epoch 199/200
Training loss for epoch 199: 1.0091499090194702
Training accuracy for epoch 199: 0.7756720185279846
Epoch 200/200
Training loss for epoch 200: 1.0069268941879272
Training accuracy for epoch 200: 0.7760361433029175

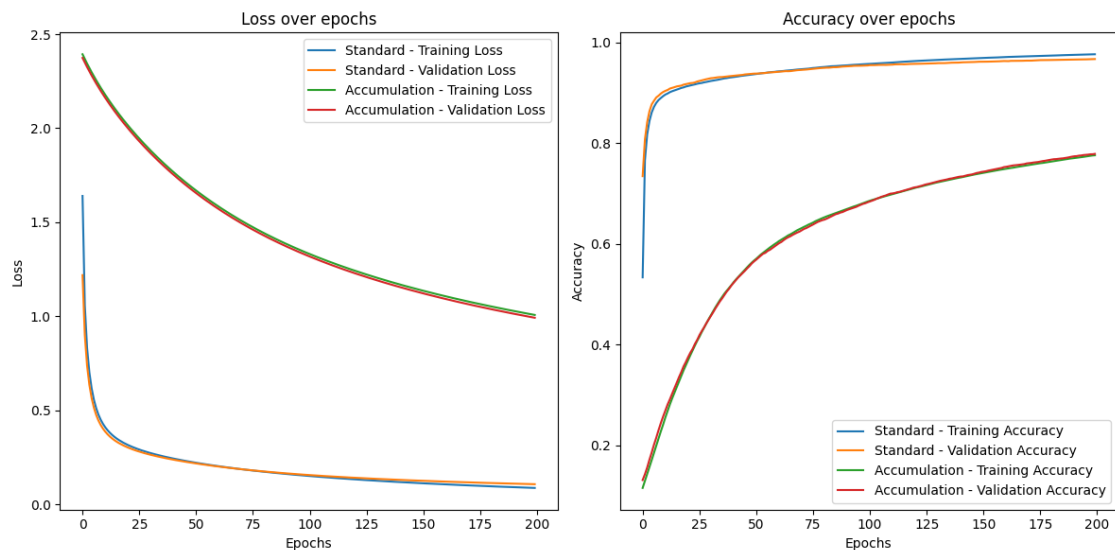
6 Visualize Results & Compare Different Approaches

```
[ ]: # Plot training and validation loss over each epoch
plt.figure(figsize=(12, 6))

plt.subplot(1, 2, 1)
plt.plot(train_losses_standard, label='Standard - Training Loss')
plt.plot(val_losses_standard, label='Standard - Validation Loss')
plt.plot(train_losses, label='Accumulation - Training Loss')
plt.plot(val_losses, label='Accumulation - Validation Loss')
plt.title('Loss over epochs')
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.legend()

# Plot training and validation accuracy over each epoch
plt.subplot(1, 2, 2)
plt.plot(train_accuracies_standard, label='Standard - Training Accuracy')
plt.plot(val_accuracies_standard, label='Standard - Validation Accuracy')
plt.plot(train_accuracies, label='Accumulation - Training Accuracy')
plt.plot(val_accuracies, label='Accumulation - Validation Accuracy')
plt.title('Accuracy over epochs')
plt.xlabel('Epochs')
plt.ylabel('Accuracy')
plt.legend()

plt.tight_layout()
plt.show()
```



7 Confusion Matrix & Classification Report

```
[ ]: from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
import seaborn as sns

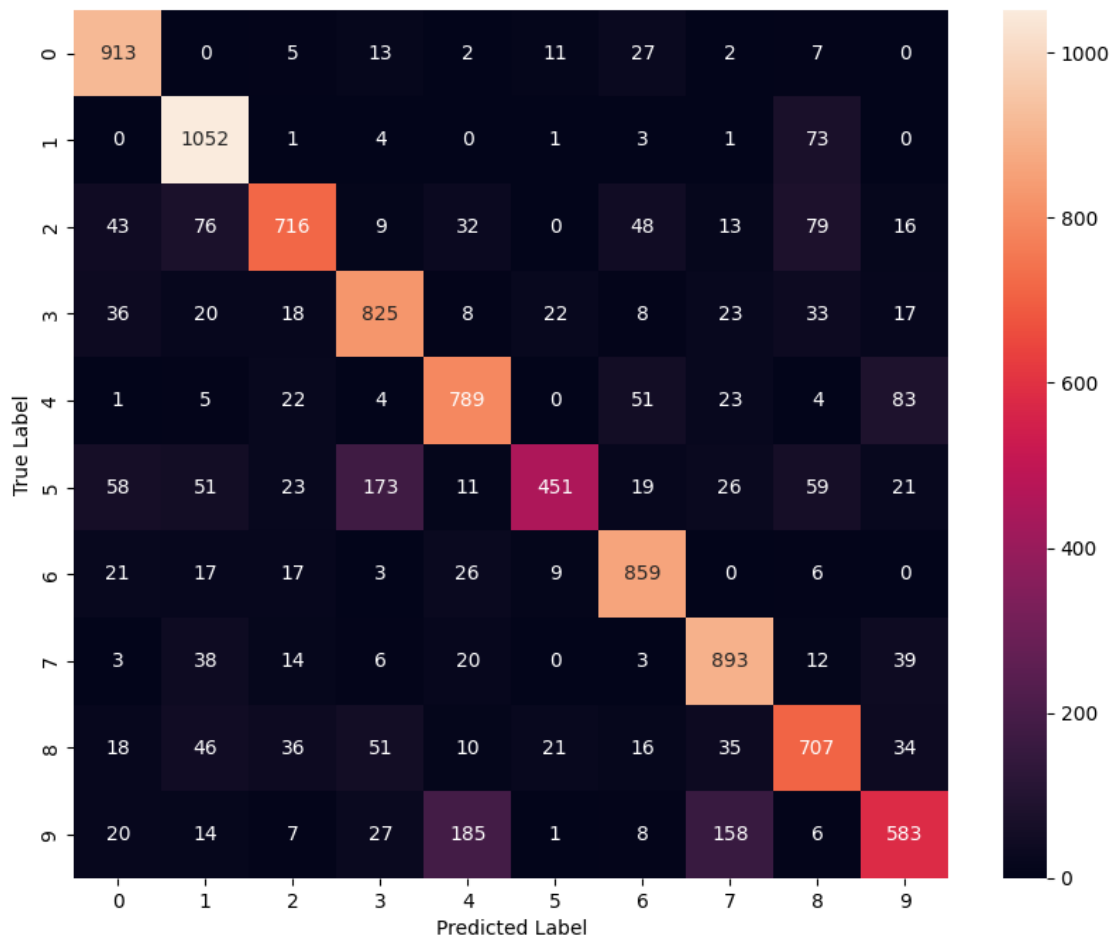
# Predict the values from the validation dataset
y_pred = a_model.predict(X_test_mnist)
# Convert predictions classes to one hot vectors
y_pred_classes = np.argmax(y_pred, axis = 1)

# compute the confusion matrix
confusion_mtx = confusion_matrix(y_test_mnist, y_pred_classes)

# plot the confusion matrix
plt.figure(figsize=(10,8))
sns.heatmap(confusion_mtx, annot=True, fmt='d')
plt.ylabel('True Label')
plt.xlabel('Predicted Label')
plt.show()

# Classification report
print(classification_report(y_test_mnist, y_pred_classes))
```

313/313 [=====] - 1s 2ms/step



	precision	recall	f1-score	support
0	0.82	0.93	0.87	980
1	0.80	0.93	0.86	1135
2	0.83	0.69	0.76	1032
3	0.74	0.82	0.78	1010
4	0.73	0.80	0.76	982
5	0.87	0.51	0.64	892
6	0.82	0.90	0.86	958
7	0.76	0.87	0.81	1028
8	0.72	0.73	0.72	974
9	0.74	0.58	0.65	1009
accuracy			0.78	10000
macro avg	0.78	0.77	0.77	10000
weighted avg	0.78	0.78	0.77	10000

8 Fashion MNIST Version

```
[ ]: # First we need to create a new model with the same starting weights

# Set the seed
tf.random.set_seed(42)

# Create the standard model
s2_model = create_model()

# Create the accumulation model
a2_model = create_model()

# We will use the same optimizer and loss function as before

s2_model.compile(optimizer=optimizer, loss=loss_fn, metrics=['accuracy'])

a2_model.compile(optimizer=optimizer, loss=loss_fn, metrics=['accuracy'])

[ ]: # Initialize lists to store metrics
train_losses_standard_fashion = []
val_losses_standard_fashion = []
train_accuracies_standard_fashion = []
val_accuracies_standard_fashion = []

# Number of epochs to train for
num_epochs = 200

for epoch in range(num_epochs):
    print(f"Epoch {epoch+1}/{num_epochs}")

    # Initialize lists to store metrics for each batch
    batch_losses = []
    batch_accuracies = []

    # Iterate over the batches of the dataset
    for step, (X_batch, y_batch) in enumerate(data_fashion):
        with tf.GradientTape() as tape:
            # Forward pass
            y_pred = s2_model(X_batch, training=True)
            # Compute the loss value
            loss_value = loss_fn(y_batch, y_pred)

            # Compute gradients
            grads = tape.gradient(loss_value, s2_model.trainable_weights)

            # Update weights
```

```

optimizer.apply_gradients(zip(grads, s2_model.trainable_weights))

# Log metrics at the end of each epoch
if step == len(data_fashion) - 1:
    train_loss = np.mean(batch_losses)
    train_accuracy = np.mean(batch_accuracies)
    print(f"Training loss for epoch {epoch+1}: {train_loss}")
    print(f"Training accuracy for epoch {epoch+1}: {train_accuracy}")

# Store metrics for this batch
batch_losses.append(loss_value.numpy())
batch_accuracies.append(tf.reduce_mean(tf.cast(tf.equal(tf.
↪cast(y_batch, tf.int64), tf.argmax(y_pred, axis=1)), tf.float32)).numpy())

# Compute average metrics for this epoch
train_losses_standard_fashion.append(np.mean(batch_losses))
train_accuracies_standard_fashion.append(np.mean(batch_accuracies))

# Compute validation metrics for this epoch
y_val_pred = s2_model(X_test_fashion, training=False)
val_loss = np.mean(loss_fn(y_test_fashion, y_val_pred).numpy())
val_accuracy = np.mean(tf.cast(tf.equal(y_test_fashion, tf.
↪argmax(y_val_pred, axis=1)), tf.float32).numpy())

val_losses_standard_fashion.append(val_loss)
val_accuracies_standard_fashion.append(val_accuracy)

```

Epoch 1/200

Training loss for epoch 1: 1.6180275678634644

Training accuracy for epoch 1: 0.47863247990608215

Epoch 2/200

Training loss for epoch 2: 1.1347659826278687

Training accuracy for epoch 2: 0.674195408821106

Epoch 3/200

Training loss for epoch 3: 0.9393534064292908

Training accuracy for epoch 3: 0.7248430848121643

Epoch 4/200

Training loss for epoch 4: 0.8261443972587585

Training accuracy for epoch 4: 0.7475460767745972

Epoch 5/200

Training loss for epoch 5: 0.753303587436676

Training accuracy for epoch 5: 0.7630375027656555

Epoch 6/200

Training loss for epoch 6: 0.7021108269691467

Training accuracy for epoch 6: 0.7747395634651184

Epoch 7/200

Training loss for epoch 7: 0.6635023951530457

Training accuracy for epoch 7: 0.7841880321502686
Epoch 8/200
Training loss for epoch 8: 0.6328722834587097
Training accuracy for epoch 8: 0.7921174168586731
Epoch 9/200
Training loss for epoch 9: 0.6077283620834351
Training accuracy for epoch 9: 0.7995793223381042
Epoch 10/200
Training loss for epoch 10: 0.586616575717926
Training accuracy for epoch 10: 0.8052717447280884
Epoch 11/200
Training loss for epoch 11: 0.5686137080192566
Training accuracy for epoch 11: 0.810363233089447
Epoch 12/200
Training loss for epoch 12: 0.5530855059623718
Training accuracy for epoch 12: 0.8137519955635071
Epoch 13/200
Training loss for epoch 13: 0.5395700931549072
Training accuracy for epoch 13: 0.8174412250518799
Epoch 14/200
Training loss for epoch 14: 0.5277161598205566
Training accuracy for epoch 14: 0.820128858089447
Epoch 15/200
Training loss for epoch 15: 0.51724773645401
Training accuracy for epoch 15: 0.8235176205635071
Epoch 16/200
Training loss for epoch 16: 0.5079423785209656
Training accuracy for epoch 16: 0.8263221383094788
Epoch 17/200
Training loss for epoch 17: 0.49961769580841064
Training accuracy for epoch 17: 0.8290431499481201
Epoch 18/200
Training loss for epoch 18: 0.492122620344162
Training accuracy for epoch 18: 0.831697404384613
Epoch 19/200
Training loss for epoch 19: 0.4853314757347107
Training accuracy for epoch 19: 0.8334001302719116
Epoch 20/200
Training loss for epoch 20: 0.4791398048400879
Training accuracy for epoch 20: 0.8352363705635071
Epoch 21/200
Training loss for epoch 21: 0.47346076369285583
Training accuracy for epoch 21: 0.8374899625778198
Epoch 22/200
Training loss for epoch 22: 0.46822211146354675
Training accuracy for epoch 22: 0.8389089107513428
Epoch 23/200
Training loss for epoch 23: 0.4633639454841614

Training accuracy for epoch 23: 0.8403946161270142
Epoch 24/200
Training loss for epoch 24: 0.45883631706237793
Training accuracy for epoch 24: 0.8416633009910583
Epoch 25/200
Training loss for epoch 25: 0.45459774136543274
Training accuracy for epoch 25: 0.8426482081413269
Epoch 26/200
Training loss for epoch 26: 0.4506133496761322
Training accuracy for epoch 26: 0.8438000679016113
Epoch 27/200
Training loss for epoch 27: 0.44685399532318115
Training accuracy for epoch 27: 0.8447015285491943
Epoch 28/200
Training loss for epoch 28: 0.443294882774353
Training accuracy for epoch 28: 0.8457531929016113
Epoch 29/200
Training loss for epoch 29: 0.43991538882255554
Training accuracy for epoch 29: 0.8467214107513428
Epoch 30/200
Training loss for epoch 30: 0.43669748306274414
Training accuracy for epoch 30: 0.8475394248962402
Epoch 31/200
Training loss for epoch 31: 0.4336259961128235
Training accuracy for epoch 31: 0.8484241366386414
Epoch 32/200
Training loss for epoch 32: 0.4306877553462982
Training accuracy for epoch 32: 0.8494925498962402
Epoch 33/200
Training loss for epoch 33: 0.4278714060783386
Training accuracy for epoch 33: 0.8505275249481201
Epoch 34/200
Training loss for epoch 34: 0.4251670241355896
Training accuracy for epoch 34: 0.8514456748962402
Epoch 35/200
Training loss for epoch 35: 0.4225657880306244
Training accuracy for epoch 35: 0.8523470759391785
Epoch 36/200
Training loss for epoch 36: 0.42006003856658936
Training accuracy for epoch 36: 0.8533820509910583
Epoch 37/200
Training loss for epoch 37: 0.41764283180236816
Training accuracy for epoch 37: 0.8541666865348816
Epoch 38/200
Training loss for epoch 38: 0.4153082072734833
Training accuracy for epoch 38: 0.8550180196762085
Epoch 39/200
Training loss for epoch 39: 0.4130505621433258

Training accuracy for epoch 39: 0.8554186820983887
Epoch 40/200
Training loss for epoch 40: 0.41086480021476746
Training accuracy for epoch 40: 0.8560529947280884
Epoch 41/200
Training loss for epoch 41: 0.408746600151062
Training accuracy for epoch 41: 0.856887698173523
Epoch 42/200
Training loss for epoch 42: 0.4066917598247528
Training accuracy for epoch 42: 0.8574719429016113
Epoch 43/200
Training loss for epoch 43: 0.40469661355018616
Training accuracy for epoch 43: 0.8581563830375671
Epoch 44/200
Training loss for epoch 44: 0.4027577042579651
Training accuracy for epoch 44: 0.8587072491645813
Epoch 45/200
Training loss for epoch 45: 0.40087181329727173
Training accuracy for epoch 45: 0.8593583106994629
Epoch 46/200
Training loss for epoch 46: 0.39903607964515686
Training accuracy for epoch 46: 0.8601095080375671
Epoch 47/200
Training loss for epoch 47: 0.397247850894928
Training accuracy for epoch 47: 0.8606770634651184
Epoch 48/200
Training loss for epoch 48: 0.3955046236515045
Training accuracy for epoch 48: 0.8612279891967773
Epoch 49/200
Training loss for epoch 49: 0.39380407333374023
Training accuracy for epoch 49: 0.861695408821106
Epoch 50/200
Training loss for epoch 50: 0.39214420318603516
Training accuracy for epoch 50: 0.8622796535491943
Epoch 51/200
Training loss for epoch 51: 0.39052292704582214
Training accuracy for epoch 51: 0.8628138303756714
Epoch 52/200
Training loss for epoch 52: 0.3889385163784027
Training accuracy for epoch 52: 0.8630809187889099
Epoch 53/200
Training loss for epoch 53: 0.38738903403282166
Training accuracy for epoch 53: 0.8635316491127014
Epoch 54/200
Training loss for epoch 54: 0.385873019695282
Training accuracy for epoch 54: 0.8640324473381042
Epoch 55/200
Training loss for epoch 55: 0.3843890130519867

Training accuracy for epoch 55: 0.8644664883613586
Epoch 56/200
Training loss for epoch 56: 0.38293564319610596
Training accuracy for epoch 56: 0.8648504018783569
Epoch 57/200
Training loss for epoch 57: 0.38151153922080994
Training accuracy for epoch 57: 0.8651843070983887
Epoch 58/200
Training loss for epoch 58: 0.38011544942855835
Training accuracy for epoch 58: 0.8656517267227173
Epoch 59/200
Training loss for epoch 59: 0.37874624133110046
Training accuracy for epoch 59: 0.866135835647583
Epoch 60/200
Training loss for epoch 60: 0.3774029314517975
Training accuracy for epoch 60: 0.8665865659713745
Epoch 61/200
Training loss for epoch 61: 0.37608441710472107
Training accuracy for epoch 61: 0.867003858089447
Epoch 62/200
Training loss for epoch 62: 0.37478965520858765
Training accuracy for epoch 62: 0.8672041893005371
Epoch 63/200
Training loss for epoch 63: 0.37351804971694946
Training accuracy for epoch 63: 0.867371141910553
Epoch 64/200
Training loss for epoch 64: 0.37226852774620056
Training accuracy for epoch 64: 0.86790531873703
Epoch 65/200
Training loss for epoch 65: 0.3710402846336365
Training accuracy for epoch 65: 0.8682892918586731
Epoch 66/200
Training loss for epoch 66: 0.3698325753211975
Training accuracy for epoch 66: 0.8688234686851501
Epoch 67/200
Training loss for epoch 67: 0.3686447739601135
Training accuracy for epoch 67: 0.8692073822021484
Epoch 68/200
Training loss for epoch 68: 0.367476224899292
Training accuracy for epoch 68: 0.8696581125259399
Epoch 69/200
Training loss for epoch 69: 0.36632612347602844
Training accuracy for epoch 69: 0.8701088428497314
Epoch 70/200
Training loss for epoch 70: 0.3651939034461975
Training accuracy for epoch 70: 0.8704093098640442
Epoch 71/200
Training loss for epoch 71: 0.36407917737960815

Training accuracy for epoch 71: 0.8707097768783569
Epoch 72/200
Training loss for epoch 72: 0.3629811704158783
Training accuracy for epoch 72: 0.8709769248962402
Epoch 73/200
Training loss for epoch 73: 0.36189940571784973
Training accuracy for epoch 73: 0.8715111017227173
Epoch 74/200
Training loss for epoch 74: 0.3608335256576538
Training accuracy for epoch 74: 0.8720786571502686
Epoch 75/200
Training loss for epoch 75: 0.3597828149795532
Training accuracy for epoch 75: 0.8723791241645813
Epoch 76/200
Training loss for epoch 76: 0.3587470054626465
Training accuracy for epoch 76: 0.8726462125778198
Epoch 77/200
Training loss for epoch 77: 0.3577255606651306
Training accuracy for epoch 77: 0.8729467391967773
Epoch 78/200
Training loss for epoch 78: 0.3567180931568146
Training accuracy for epoch 78: 0.8732638955116272
Epoch 79/200
Training loss for epoch 79: 0.3557242453098297
Training accuracy for epoch 79: 0.8735977411270142
Epoch 80/200
Training loss for epoch 80: 0.3547435402870178
Training accuracy for epoch 80: 0.8739817142486572
Epoch 81/200
Training loss for epoch 81: 0.35377562046051025
Training accuracy for epoch 81: 0.8744825124740601
Epoch 82/200
Training loss for epoch 82: 0.35282018780708313
Training accuracy for epoch 82: 0.8747662901878357
Epoch 83/200
Training loss for epoch 83: 0.3518768846988678
Training accuracy for epoch 83: 0.8750333786010742
Epoch 84/200
Training loss for epoch 84: 0.3509453535079956
Training accuracy for epoch 84: 0.8754841089248657
Epoch 85/200
Training loss for epoch 85: 0.3500252962112427
Training accuracy for epoch 85: 0.8757678866386414
Epoch 86/200
Training loss for epoch 86: 0.34911632537841797
Training accuracy for epoch 86: 0.8761017918586731
Epoch 87/200
Training loss for epoch 87: 0.3482183516025543

Training accuracy for epoch 87: 0.8763187527656555
Epoch 88/200
Training loss for epoch 88: 0.34733086824417114
Training accuracy for epoch 88: 0.876585841178894
Epoch 89/200
Training loss for epoch 89: 0.3464536666870117
Training accuracy for epoch 89: 0.8768696784973145
Epoch 90/200
Training loss for epoch 90: 0.34558653831481934
Training accuracy for epoch 90: 0.8772035241127014
Epoch 91/200
Training loss for epoch 91: 0.3447292149066925
Training accuracy for epoch 91: 0.8774539232254028
Epoch 92/200
Training loss for epoch 92: 0.34388136863708496
Training accuracy for epoch 92: 0.8776542544364929
Epoch 93/200
Training loss for epoch 93: 0.34304285049438477
Training accuracy for epoch 93: 0.8779213428497314
Epoch 94/200
Training loss for epoch 94: 0.34221339225769043
Training accuracy for epoch 94: 0.8780214786529541
Epoch 95/200
Training loss for epoch 95: 0.34139272570610046
Training accuracy for epoch 95: 0.8783220052719116
Epoch 96/200
Training loss for epoch 96: 0.34058070182800293
Training accuracy for epoch 96: 0.8788895606994629
Epoch 97/200
Training loss for epoch 97: 0.3397771120071411
Training accuracy for epoch 97: 0.8793402910232544
Epoch 98/200
Training loss for epoch 98: 0.33898165822029114
Training accuracy for epoch 98: 0.8796073794364929
Epoch 99/200
Training loss for epoch 99: 0.3381943106651306
Training accuracy for epoch 99: 0.8798243999481201
Epoch 100/200
Training loss for epoch 100: 0.3374147117137909
Training accuracy for epoch 100: 0.8801081776618958
Epoch 101/200
Training loss for epoch 101: 0.3366427421569824
Training accuracy for epoch 101: 0.8803418874740601
Epoch 102/200
Training loss for epoch 102: 0.33587825298309326
Training accuracy for epoch 102: 0.880525529384613
Epoch 103/200
Training loss for epoch 103: 0.3351210951805115

Training accuracy for epoch 103: 0.880675733089447
Epoch 104/200
Training loss for epoch 104: 0.3343709707260132
Training accuracy for epoch 104: 0.8809762001037598
Epoch 105/200
Training loss for epoch 105: 0.333627849817276
Training accuracy for epoch 105: 0.881226658821106
Epoch 106/200
Training loss for epoch 106: 0.3328914940357208
Training accuracy for epoch 106: 0.8815104365348816
Epoch 107/200
Training loss for epoch 107: 0.3321617841720581
Training accuracy for epoch 107: 0.8817107081413269
Epoch 108/200
Training loss for epoch 108: 0.3314385712146759
Training accuracy for epoch 108: 0.8819778561592102
Epoch 109/200
Training loss for epoch 109: 0.3307216763496399
Training accuracy for epoch 109: 0.88209468126297
Epoch 110/200
Training loss for epoch 110: 0.3300110697746277
Training accuracy for epoch 110: 0.8824619650840759
Epoch 111/200
Training loss for epoch 111: 0.3293064832687378
Training accuracy for epoch 111: 0.8827123641967773
Epoch 112/200
Training loss for epoch 112: 0.32860785722732544
Training accuracy for epoch 112: 0.8827958106994629
Epoch 113/200
Training loss for epoch 113: 0.3279150724411011
Training accuracy for epoch 113: 0.8833299875259399
Epoch 114/200
Training loss for epoch 114: 0.3272278606891632
Training accuracy for epoch 114: 0.8835636973381042
Epoch 115/200
Training loss for epoch 115: 0.32654625177383423
Training accuracy for epoch 115: 0.8837306499481201
Epoch 116/200
Training loss for epoch 116: 0.3258701264858246
Training accuracy for epoch 116: 0.88404780626297
Epoch 117/200
Training loss for epoch 117: 0.32519927620887756
Training accuracy for epoch 117: 0.8842648267745972
Epoch 118/200
Training loss for epoch 118: 0.32453370094299316
Training accuracy for epoch 118: 0.8844651579856873
Epoch 119/200
Training loss for epoch 119: 0.3238731920719147

Training accuracy for epoch 119: 0.8848490715026855
Epoch 120/200
Training loss for epoch 120: 0.3232176601886749
Training accuracy for epoch 120: 0.8850160241127014
Epoch 121/200
Training loss for epoch 121: 0.32256701588630676
Training accuracy for epoch 121: 0.8851495981216431
Epoch 122/200
Training loss for epoch 122: 0.3219212591648102
Training accuracy for epoch 122: 0.8854500651359558
Epoch 123/200
Training loss for epoch 123: 0.3212800920009613
Training accuracy for epoch 123: 0.8858006000518799
Epoch 124/200
Training loss for epoch 124: 0.3206436336040497
Training accuracy for epoch 124: 0.8860676884651184
Epoch 125/200
Training loss for epoch 125: 0.32001155614852905
Training accuracy for epoch 125: 0.8864349722862244
Epoch 126/200
Training loss for epoch 126: 0.3193839490413666
Training accuracy for epoch 126: 0.8867521286010742
Epoch 127/200
Training loss for epoch 127: 0.3187606632709503
Training accuracy for epoch 127: 0.887052595615387
Epoch 128/200
Training loss for epoch 128: 0.3181416094303131
Training accuracy for epoch 128: 0.8873698115348816
Epoch 129/200
Training loss for epoch 129: 0.31752675771713257
Training accuracy for epoch 129: 0.8875533938407898
Epoch 130/200
Training loss for epoch 130: 0.31691598892211914
Training accuracy for epoch 130: 0.8877036571502686
Epoch 131/200
Training loss for epoch 131: 0.3163091540336609
Training accuracy for epoch 131: 0.8879707455635071
Epoch 132/200
Training loss for epoch 132: 0.31570619344711304
Training accuracy for epoch 132: 0.8882211446762085
Epoch 133/200
Training loss for epoch 133: 0.315107136964798
Training accuracy for epoch 133: 0.8884047865867615
Epoch 134/200
Training loss for epoch 134: 0.31451186537742615
Training accuracy for epoch 134: 0.8886218070983887
Epoch 135/200
Training loss for epoch 135: 0.3139203190803528

Training accuracy for epoch 135: 0.8887887001037598
Epoch 136/200
Training loss for epoch 136: 0.31333237886428833
Training accuracy for epoch 136: 0.889005720615387
Epoch 137/200
Training loss for epoch 137: 0.3127480745315552
Training accuracy for epoch 137: 0.8892561197280884
Epoch 138/200
Training loss for epoch 138: 0.31216713786125183
Training accuracy for epoch 138: 0.8895900249481201
Epoch 139/200
Training loss for epoch 139: 0.31158971786499023
Training accuracy for epoch 139: 0.8897736668586731
Epoch 140/200
Training loss for epoch 140: 0.31101569533348083
Training accuracy for epoch 140: 0.8899238705635071
Epoch 141/200
Training loss for epoch 141: 0.3104449212551117
Training accuracy for epoch 141: 0.8901575803756714
Epoch 142/200
Training loss for epoch 142: 0.30987751483917236
Training accuracy for epoch 142: 0.8904246687889099
Epoch 143/200
Training loss for epoch 143: 0.30931323766708374
Training accuracy for epoch 143: 0.8907919526100159
Epoch 144/200
Training loss for epoch 144: 0.308752179145813
Training accuracy for epoch 144: 0.8909254670143127
Epoch 145/200
Training loss for epoch 145: 0.3081941306591034
Training accuracy for epoch 145: 0.8910423517227173
Epoch 146/200
Training loss for epoch 146: 0.30763915181159973
Training accuracy for epoch 146: 0.8912426829338074
Epoch 147/200
Training loss for epoch 147: 0.3070872128009796
Training accuracy for epoch 147: 0.8914095759391785
Epoch 148/200
Training loss for epoch 148: 0.3065381944179535
Training accuracy for epoch 148: 0.8915598392486572
Epoch 149/200
Training loss for epoch 149: 0.30599209666252136
Training accuracy for epoch 149: 0.8916933536529541
Epoch 150/200
Training loss for epoch 150: 0.3054488003253937
Training accuracy for epoch 150: 0.8918269276618958
Epoch 151/200
Training loss for epoch 151: 0.3049083650112152

Training accuracy for epoch 151: 0.8921273946762085
Epoch 152/200
Training loss for epoch 152: 0.3043707311153412
Training accuracy for epoch 152: 0.8923444151878357
Epoch 153/200
Training loss for epoch 153: 0.30383574962615967
Training accuracy for epoch 153: 0.8925948143005371
Epoch 154/200
Training loss for epoch 154: 0.3033035099506378
Training accuracy for epoch 154: 0.8927951455116272
Epoch 155/200
Training loss for epoch 155: 0.30277395248413086
Training accuracy for epoch 155: 0.8928452134132385
Epoch 156/200
Training loss for epoch 156: 0.30224692821502686
Training accuracy for epoch 156: 0.8930622339248657
Epoch 157/200
Training loss for epoch 157: 0.3017224967479706
Training accuracy for epoch 157: 0.8932124972343445
Epoch 158/200
Training loss for epoch 158: 0.30120059847831726
Training accuracy for epoch 158: 0.8933960795402527
Epoch 159/200
Training loss for epoch 159: 0.3006812632083893
Training accuracy for epoch 159: 0.8936131000518799
Epoch 160/200
Training loss for epoch 160: 0.3001643121242523
Training accuracy for epoch 160: 0.8936966061592102
Epoch 161/200
Training loss for epoch 161: 0.2996498644351959
Training accuracy for epoch 161: 0.8939303159713745
Epoch 162/200
Training loss for epoch 162: 0.29913774132728577
Training accuracy for epoch 162: 0.8940471410751343
Epoch 163/200
Training loss for epoch 163: 0.2986280024051666
Training accuracy for epoch 163: 0.8940638303756714
Epoch 164/200
Training loss for epoch 164: 0.2981206178665161
Training accuracy for epoch 164: 0.8942140936851501
Epoch 165/200
Training loss for epoch 165: 0.29761558771133423
Training accuracy for epoch 165: 0.894347608089447
Epoch 166/200
Training loss for epoch 166: 0.2971127927303314
Training accuracy for epoch 166: 0.8945145606994629
Epoch 167/200
Training loss for epoch 167: 0.2966122329235077

Training accuracy for epoch 167: 0.8947482705116272
Epoch 168/200
Training loss for epoch 168: 0.29611387848854065
Training accuracy for epoch 168: 0.8950153589248657
Epoch 169/200
Training loss for epoch 169: 0.2956177592277527
Training accuracy for epoch 169: 0.8951823115348816
Epoch 170/200
Training loss for epoch 170: 0.29512378573417664
Training accuracy for epoch 170: 0.8954660892486572
Epoch 171/200
Training loss for epoch 171: 0.2946319580078125
Training accuracy for epoch 171: 0.8957331776618958
Epoch 172/200
Training loss for epoch 172: 0.2941422462463379
Training accuracy for epoch 172: 0.8959001302719116
Epoch 173/200
Training loss for epoch 173: 0.2936546504497528
Training accuracy for epoch 173: 0.8962172865867615
Epoch 174/200
Training loss for epoch 174: 0.2931691110134125
Training accuracy for epoch 174: 0.8963675498962402
Epoch 175/200
Training loss for epoch 175: 0.29268550872802734
Training accuracy for epoch 175: 0.8966179490089417
Epoch 176/200
Training loss for epoch 176: 0.29220402240753174
Training accuracy for epoch 176: 0.8969017267227173
Epoch 177/200
Training loss for epoch 177: 0.2917245328426361
Training accuracy for epoch 177: 0.8971354365348816
Epoch 178/200
Training loss for epoch 178: 0.2912469506263733
Training accuracy for epoch 178: 0.8972689509391785
Epoch 179/200
Training loss for epoch 179: 0.29077136516571045
Training accuracy for epoch 179: 0.897385835647583
Epoch 180/200
Training loss for epoch 180: 0.29029765725135803
Training accuracy for epoch 180: 0.897536039352417
Epoch 181/200
Training loss for epoch 181: 0.2898259162902832
Training accuracy for epoch 181: 0.89771968126297
Epoch 182/200
Training loss for epoch 182: 0.289355993270874
Training accuracy for epoch 182: 0.8979867696762085
Epoch 183/200
Training loss for epoch 183: 0.28888803720474243

Training accuracy for epoch 183: 0.8980368375778198
Epoch 184/200
Training loss for epoch 184: 0.28842177987098694
Training accuracy for epoch 184: 0.8981537222862244
Epoch 185/200
Training loss for epoch 185: 0.28795742988586426
Training accuracy for epoch 185: 0.8982705473899841
Epoch 186/200
Training loss for epoch 186: 0.2874948978424072
Training accuracy for epoch 186: 0.8984375
Epoch 187/200
Training loss for epoch 187: 0.28703415393829346
Training accuracy for epoch 187: 0.8985543251037598
Epoch 188/200
Training loss for epoch 188: 0.2865751087665558
Training accuracy for epoch 188: 0.8986712098121643
Epoch 189/200
Training loss for epoch 189: 0.2861178517341614
Training accuracy for epoch 189: 0.8988214731216431
Epoch 190/200
Training loss for epoch 190: 0.2856622636318207
Training accuracy for epoch 190: 0.8989049196243286
Epoch 191/200
Training loss for epoch 191: 0.28520840406417847
Training accuracy for epoch 191: 0.8990885615348816
Epoch 192/200
Training loss for epoch 192: 0.28475630283355713
Training accuracy for epoch 192: 0.8993222713470459
Epoch 193/200
Training loss for epoch 193: 0.2843058407306671
Training accuracy for epoch 193: 0.8995058536529541
Epoch 194/200
Training loss for epoch 194: 0.28385695815086365
Training accuracy for epoch 194: 0.8996060490608215
Epoch 195/200
Training loss for epoch 195: 0.28340980410575867
Training accuracy for epoch 195: 0.8997562527656555
Epoch 196/200
Training loss for epoch 196: 0.282964289188385
Training accuracy for epoch 196: 0.8998731374740601
Epoch 197/200
Training loss for epoch 197: 0.2825203239917755
Training accuracy for epoch 197: 0.8999232053756714
Epoch 198/200
Training loss for epoch 198: 0.28207793831825256
Training accuracy for epoch 198: 0.9000901579856873
Epoch 199/200
Training loss for epoch 199: 0.28163716197013855

Training accuracy for epoch 199: 0.9001902937889099
Epoch 200/200
Training loss for epoch 200: 0.2811979651451111
Training accuracy for epoch 200: 0.9002904891967773

9 Accumulation Model

```
[ ]: # Initialize lists to store metrics
a_train_losses_fashion = []
a_val_losses_fashion = []
a_train_accuracies_fashion = []
a_val_accuracies_fashion = []

# Number of epochs to train for
num_epochs = 200

# Initialize accumulators
accumulated_grads = [tf.zeros_like(w) for w in a2_model.trainable_weights]

for epoch in range(num_epochs):
    print(f"Epoch {epoch+1}/{num_epochs}")

    # Initialize lists to store metrics for each batch
    batch_losses = []
    batch_accuracies = []

    # Iterate over the batches of the dataset
    for step, (X_batch, y_batch) in enumerate(data_fashion):
        with tf.GradientTape() as tape:
            # Forward pass
            y_pred = a2_model(X_batch, training=True)
            # Compute the loss value
            loss_value = loss_fn(y_batch, y_pred)

            # Compute gradients - backward pass
            grads = tape.gradient(loss_value, a2_model.trainable_weights)

            # If it's the final batch, adjust the gradients
            if (step + 1) == len(data_fashion):
                final_batch_size = len(X_batch)
                grads = [g * final_batch_size / batch_size for g in grads]

            # Accumulate gradients
            accumulated_grads = [acc_g + g for acc_g, g in zip(accumulated_grads,
↪grads)]

    # Store metrics for this batch
```

```

        batch_losses.append(loss_value.numpy())
        batch_accuracies.append(tf.reduce_mean(tf.cast(tf.equal(tf.
↪cast(y_batch, tf.int64), tf.argmax(y_pred, axis=1)), tf.float32)))

        # Normalize the accumulated gradients
        accumulated_grads = [acc_g / len(data_fashion) for acc_g in
↪accumulated_grads]

        # Update weights after all batches are processed
        optimizer.apply_gradients(zip(accumulated_grads, a2_model.
↪trainable_weights))

        # Compute average metrics for this epoch
        avg_train_loss = np.mean(batch_losses)
        avg_train_accuracy = np.mean(batch_accuracies)

        # Print the loss and accuracy for each epoch
        print(f"Training loss for epoch {epoch+1}: {avg_train_loss}")
        print(f"Training accuracy for epoch {epoch+1}: {avg_train_accuracy}")

        # Append average metrics for this epoch
        a_train_losses_fashion.append(avg_train_loss)
        a_train_accuracies_fashion.append(avg_train_accuracy)

        # Compute validation metrics for this epoch
        y_val_pred = a2_model(X_test_fashion, training=False)
        val_loss = np.mean(loss_fn(y_test_fashion, y_val_pred).numpy())
        val_accuracy = np.mean(tf.cast(tf.equal(y_test_fashion, tf.
↪argmax(y_val_pred, axis=1)), tf.float32).numpy())

        a_val_losses_fashion.append(val_loss)
        a_val_accuracies_fashion.append(val_accuracy)

        # Reset the accumulators for the next epoch
        accumulated_grads = [tf.zeros_like(w) for w in a2_model.trainable_weights]

```

Epoch 1/200

Training loss for epoch 1: 2.310701608657837

Training accuracy for epoch 1: 0.0833829864859581

Epoch 2/200

Training loss for epoch 2: 2.2763760089874268

Training accuracy for epoch 2: 0.08476783335208893

Epoch 3/200

Training loss for epoch 3: 2.2461938858032227

Training accuracy for epoch 3: 0.09358447790145874

Epoch 4/200

Training loss for epoch 4: 2.218925952911377

Training accuracy for epoch 4: 0.10981086641550064
Epoch 5/200
Training loss for epoch 5: 2.1937613487243652
Training accuracy for epoch 5: 0.13298353552818298
Epoch 6/200
Training loss for epoch 6: 2.170168161392212
Training accuracy for epoch 6: 0.15806519985198975
Epoch 7/200
Training loss for epoch 7: 2.147792339324951
Training accuracy for epoch 7: 0.18788620829582214
Epoch 8/200
Training loss for epoch 8: 2.1263937950134277
Training accuracy for epoch 8: 0.21815964579582214
Epoch 9/200
Training loss for epoch 9: 2.105804920196533
Training accuracy for epoch 9: 0.2469213455915451
Epoch 10/200
Training loss for epoch 10: 2.085905075073242
Training accuracy for epoch 10: 0.2751975357532501
Epoch 11/200
Training loss for epoch 11: 2.06660532951355
Training accuracy for epoch 11: 0.3026902377605438
Epoch 12/200
Training loss for epoch 12: 2.0478382110595703
Training accuracy for epoch 12: 0.32658568024635315
Epoch 13/200
Training loss for epoch 13: 2.0295519828796387
Training accuracy for epoch 13: 0.34841763973236084
Epoch 14/200
Training loss for epoch 14: 2.011705160140991
Training accuracy for epoch 14: 0.3671819865703583
Epoch 15/200
Training loss for epoch 15: 1.9942667484283447
Training accuracy for epoch 15: 0.3846718370914459
Epoch 16/200
Training loss for epoch 16: 1.977210521697998
Training accuracy for epoch 16: 0.3997892141342163
Epoch 17/200
Training loss for epoch 17: 1.9605166912078857
Training accuracy for epoch 17: 0.4132404327392578
Epoch 18/200
Training loss for epoch 18: 1.9441691637039185
Training accuracy for epoch 18: 0.4256874620914459
Epoch 19/200
Training loss for epoch 19: 1.9281560182571411
Training accuracy for epoch 19: 0.4376268982887268
Epoch 20/200
Training loss for epoch 20: 1.9124668836593628

Training accuracy for epoch 20: 0.4489760100841522
Epoch 21/200
Training loss for epoch 21: 1.8970942497253418
Training accuracy for epoch 21: 0.46087127923965454
Epoch 22/200
Training loss for epoch 22: 1.8820316791534424
Training accuracy for epoch 22: 0.4707527756690979
Epoch 23/200
Training loss for epoch 23: 1.8672746419906616
Training accuracy for epoch 23: 0.47960805892944336
Epoch 24/200
Training loss for epoch 24: 1.8528181314468384
Training accuracy for epoch 24: 0.4886784851551056
Epoch 25/200
Training loss for epoch 25: 1.838659405708313
Training accuracy for epoch 25: 0.49676135182380676
Epoch 26/200
Training loss for epoch 26: 1.8247935771942139
Training accuracy for epoch 26: 0.5047283172607422
Epoch 27/200
Training loss for epoch 27: 1.811218023300171
Training accuracy for epoch 27: 0.5115587711334229
Epoch 28/200
Training loss for epoch 28: 1.797929286956787
Training accuracy for epoch 28: 0.5182843208312988
Epoch 29/200
Training loss for epoch 29: 1.7849234342575073
Training accuracy for epoch 29: 0.5249713063240051
Epoch 30/200
Training loss for epoch 30: 1.7721965312957764
Training accuracy for epoch 30: 0.5304279327392578
Epoch 31/200
Training loss for epoch 31: 1.75974440574646
Training accuracy for epoch 31: 0.5357962250709534
Epoch 32/200
Training loss for epoch 32: 1.747562289237976
Training accuracy for epoch 32: 0.5413632392883301
Epoch 33/200
Training loss for epoch 33: 1.7356460094451904
Training accuracy for epoch 33: 0.5470625758171082
Epoch 34/200
Training loss for epoch 34: 1.7239904403686523
Training accuracy for epoch 34: 0.5517798662185669
Epoch 35/200
Training loss for epoch 35: 1.712589979171753
Training accuracy for epoch 35: 0.5556033253669739
Epoch 36/200
Training loss for epoch 36: 1.701439380645752

Training accuracy for epoch 36: 0.5597137808799744
Epoch 37/200
Training loss for epoch 37: 1.690532922744751
Training accuracy for epoch 37: 0.5639841556549072
Epoch 38/200
Training loss for epoch 38: 1.679864525794983
Training accuracy for epoch 38: 0.5680779814720154
Epoch 39/200
Training loss for epoch 39: 1.6694285869598389
Training accuracy for epoch 39: 0.571868360042572
Epoch 40/200
Training loss for epoch 40: 1.6592190265655518
Training accuracy for epoch 40: 0.575592577457428
Epoch 41/200
Training loss for epoch 41: 1.6492292881011963
Training accuracy for epoch 41: 0.5788091421127319
Epoch 42/200
Training loss for epoch 42: 1.6394541263580322
Training accuracy for epoch 42: 0.5822243690490723
Epoch 43/200
Training loss for epoch 43: 1.6298867464065552
Training accuracy for epoch 43: 0.5854851007461548
Epoch 44/200
Training loss for epoch 44: 1.6205211877822876
Training accuracy for epoch 44: 0.5885140895843506
Epoch 45/200
Training loss for epoch 45: 1.6113520860671997
Training accuracy for epoch 45: 0.5915320515632629
Epoch 46/200
Training loss for epoch 46: 1.6023730039596558
Training accuracy for epoch 46: 0.5948920845985413
Epoch 47/200
Training loss for epoch 47: 1.5935784578323364
Training accuracy for epoch 47: 0.5975514650344849
Epoch 48/200
Training loss for epoch 48: 1.5849626064300537
Training accuracy for epoch 48: 0.5999183654785156
Epoch 49/200
Training loss for epoch 49: 1.5765200853347778
Training accuracy for epoch 49: 0.6021694540977478
Epoch 50/200
Training loss for epoch 50: 1.5682454109191895
Training accuracy for epoch 50: 0.604519784450531
Epoch 51/200
Training loss for epoch 51: 1.560133457183838
Training accuracy for epoch 51: 0.6069198250770569
Epoch 52/200
Training loss for epoch 52: 1.5521790981292725

Training accuracy for epoch 52: 0.6094191670417786
Epoch 53/200
Training loss for epoch 53: 1.5443775653839111
Training accuracy for epoch 53: 0.6113226413726807
Epoch 54/200
Training loss for epoch 54: 1.5367233753204346
Training accuracy for epoch 54: 0.6135405898094177
Epoch 55/200
Training loss for epoch 55: 1.5292130708694458
Training accuracy for epoch 55: 0.6158081889152527
Epoch 56/200
Training loss for epoch 56: 1.521841287612915
Training accuracy for epoch 56: 0.6181309819221497
Epoch 57/200
Training loss for epoch 57: 1.514604091644287
Training accuracy for epoch 57: 0.6198192238807678
Epoch 58/200
Training loss for epoch 58: 1.5074975490570068
Training accuracy for epoch 58: 0.6217172145843506
Epoch 59/200
Training loss for epoch 59: 1.5005168914794922
Training accuracy for epoch 59: 0.6232565641403198
Epoch 60/200
Training loss for epoch 60: 1.4936589002609253
Training accuracy for epoch 60: 0.6252759099006653
Epoch 61/200
Training loss for epoch 61: 1.486919641494751
Training accuracy for epoch 61: 0.6268317699432373
Epoch 62/200
Training loss for epoch 62: 1.4802957773208618
Training accuracy for epoch 62: 0.628271758556366
Epoch 63/200
Training loss for epoch 63: 1.4737838506698608
Training accuracy for epoch 63: 0.6296290159225464
Epoch 64/200
Training loss for epoch 64: 1.467380404472351
Training accuracy for epoch 64: 0.6313669681549072
Epoch 65/200
Training loss for epoch 65: 1.461081862449646
Training accuracy for epoch 65: 0.6327242255210876
Epoch 66/200
Training loss for epoch 66: 1.4548858404159546
Training accuracy for epoch 66: 0.6341477036476135
Epoch 67/200
Training loss for epoch 67: 1.4487887620925903
Training accuracy for epoch 67: 0.6355711817741394
Epoch 68/200
Training loss for epoch 68: 1.4427884817123413

Training accuracy for epoch 68: 0.636746346950531
Epoch 69/200
Training loss for epoch 69: 1.4368820190429688
Training accuracy for epoch 69: 0.6381753087043762
Epoch 70/200
Training loss for epoch 70: 1.4310665130615234
Training accuracy for epoch 70: 0.6394332647323608
Epoch 71/200
Training loss for epoch 71: 1.4253395795822144
Training accuracy for epoch 71: 0.6407077312469482
Epoch 72/200
Training loss for epoch 72: 1.4196988344192505
Training accuracy for epoch 72: 0.6422194838523865
Epoch 73/200
Training loss for epoch 73: 1.4141418933868408
Training accuracy for epoch 73: 0.6434939503669739
Epoch 74/200
Training loss for epoch 74: 1.4086661338806152
Training accuracy for epoch 74: 0.6446526050567627
Epoch 75/200
Training loss for epoch 75: 1.4032700061798096
Training accuracy for epoch 75: 0.6454139947891235
Epoch 76/200
Training loss for epoch 76: 1.3979512453079224
Training accuracy for epoch 76: 0.6461753845214844
Epoch 77/200
Training loss for epoch 77: 1.3927075862884521
Training accuracy for epoch 77: 0.6471409797668457
Epoch 78/200
Training loss for epoch 78: 1.3875371217727661
Training accuracy for epoch 78: 0.6480844020843506
Epoch 79/200
Training loss for epoch 79: 1.3824385404586792
Training accuracy for epoch 79: 0.6490609645843506
Epoch 80/200
Training loss for epoch 80: 1.3774092197418213
Training accuracy for epoch 80: 0.6500871777534485
Epoch 81/200
Training loss for epoch 81: 1.3724480867385864
Training accuracy for epoch 81: 0.6508982181549072
Epoch 82/200
Training loss for epoch 82: 1.3675527572631836
Training accuracy for epoch 82: 0.6518913507461548
Epoch 83/200
Training loss for epoch 83: 1.3627221584320068
Training accuracy for epoch 83: 0.652685821056366
Epoch 84/200
Training loss for epoch 84: 1.357954740524292

Training accuracy for epoch 84: 0.6537782549858093
Epoch 85/200
Training loss for epoch 85: 1.3532484769821167
Training accuracy for epoch 85: 0.6547437906265259
Epoch 86/200
Training loss for epoch 86: 1.348602294921875
Training accuracy for epoch 86: 0.6554223895072937
Epoch 87/200
Training loss for epoch 87: 1.3440150022506714
Training accuracy for epoch 87: 0.6563493013381958
Epoch 88/200
Training loss for epoch 88: 1.3394845724105835
Training accuracy for epoch 88: 0.6569783091545105
Epoch 89/200
Training loss for epoch 89: 1.335010290145874
Training accuracy for epoch 89: 0.6580210328102112
Epoch 90/200
Training loss for epoch 90: 1.3305906057357788
Training accuracy for epoch 90: 0.6587493419647217
Epoch 91/200
Training loss for epoch 91: 1.3262242078781128
Training accuracy for epoch 91: 0.6594610810279846
Epoch 92/200
Training loss for epoch 92: 1.3219101428985596
Training accuracy for epoch 92: 0.6601396799087524
Epoch 93/200
Training loss for epoch 93: 1.3176472187042236
Training accuracy for epoch 93: 0.6607851982116699
Epoch 94/200
Training loss for epoch 94: 1.31343412399292
Training accuracy for epoch 94: 0.6615465879440308
Epoch 95/200
Training loss for epoch 95: 1.309269905090332
Training accuracy for epoch 95: 0.6622252464294434
Epoch 96/200
Training loss for epoch 96: 1.305153489112854
Training accuracy for epoch 96: 0.6630528569221497
Epoch 97/200
Training loss for epoch 97: 1.3010836839675903
Training accuracy for epoch 97: 0.663665235042572
Epoch 98/200
Training loss for epoch 98: 1.297059416770935
Training accuracy for epoch 98: 0.6642280220985413
Epoch 99/200
Training loss for epoch 99: 1.2930806875228882
Training accuracy for epoch 99: 0.6648790240287781
Epoch 100/200
Training loss for epoch 100: 1.2891452312469482

Training accuracy for epoch 100: 0.6654252409934998
Epoch 101/200
Training loss for epoch 101: 1.2852530479431152
Training accuracy for epoch 101: 0.666043221950531
Epoch 102/200
Training loss for epoch 102: 1.2814030647277832
Training accuracy for epoch 102: 0.6666169762611389
Epoch 103/200
Training loss for epoch 103: 1.2775944471359253
Training accuracy for epoch 103: 0.6674391031265259
Epoch 104/200
Training loss for epoch 104: 1.2738261222839355
Training accuracy for epoch 104: 0.6682170033454895
Epoch 105/200
Training loss for epoch 105: 1.270097255706787
Training accuracy for epoch 105: 0.6688956618309021
Epoch 106/200
Training loss for epoch 106: 1.2664076089859009
Training accuracy for epoch 106: 0.6696404814720154
Epoch 107/200
Training loss for epoch 107: 1.2627562284469604
Training accuracy for epoch 107: 0.6700873970985413
Epoch 108/200
Training loss for epoch 108: 1.2591415643692017
Training accuracy for epoch 108: 0.6705177426338196
Epoch 109/200
Training loss for epoch 109: 1.2555642127990723
Training accuracy for epoch 109: 0.6712294816970825
Epoch 110/200
Training loss for epoch 110: 1.2520227432250977
Training accuracy for epoch 110: 0.6717756986618042
Epoch 111/200
Training loss for epoch 111: 1.24851655960083
Training accuracy for epoch 111: 0.6722556948661804
Epoch 112/200
Training loss for epoch 112: 1.2450450658798218
Training accuracy for epoch 112: 0.6728184819221497
Epoch 113/200
Training loss for epoch 113: 1.2416077852249146
Training accuracy for epoch 113: 0.6733812093734741
Epoch 114/200
Training loss for epoch 114: 1.2382036447525024
Training accuracy for epoch 114: 0.6738115549087524
Epoch 115/200
Training loss for epoch 115: 1.2348326444625854
Training accuracy for epoch 115: 0.6745067238807678
Epoch 116/200
Training loss for epoch 116: 1.231493592262268

Training accuracy for epoch 116: 0.6748874187469482
Epoch 117/200
Training loss for epoch 117: 1.2281863689422607
Training accuracy for epoch 117: 0.6754336357116699
Epoch 118/200
Training loss for epoch 118: 1.2249103784561157
Training accuracy for epoch 118: 0.6760460734367371
Epoch 119/200
Training loss for epoch 119: 1.2216650247573853
Training accuracy for epoch 119: 0.6767632961273193
Epoch 120/200
Training loss for epoch 120: 1.218449592590332
Training accuracy for epoch 120: 0.6772432923316956
Epoch 121/200
Training loss for epoch 121: 1.2152637243270874
Training accuracy for epoch 121: 0.677706778049469
Epoch 122/200
Training loss for epoch 122: 1.2121074199676514
Training accuracy for epoch 122: 0.6783853769302368
Epoch 123/200
Training loss for epoch 123: 1.2089791297912598
Training accuracy for epoch 123: 0.6789371967315674
Epoch 124/200
Training loss for epoch 124: 1.2058790922164917
Training accuracy for epoch 124: 0.6794668436050415
Epoch 125/200
Training loss for epoch 125: 1.2028067111968994
Training accuracy for epoch 125: 0.6803109645843506
Epoch 126/200
Training loss for epoch 126: 1.1997617483139038
Training accuracy for epoch 126: 0.6808241009712219
Epoch 127/200
Training loss for epoch 127: 1.1967436075210571
Training accuracy for epoch 127: 0.681353747844696
Epoch 128/200
Training loss for epoch 128: 1.1937514543533325
Training accuracy for epoch 128: 0.6819165349006653
Epoch 129/200
Training loss for epoch 129: 1.19078528881073
Training accuracy for epoch 129: 0.6826117038726807
Epoch 130/200
Training loss for epoch 130: 1.18784499168396
Training accuracy for epoch 130: 0.6832406520843506
Epoch 131/200
Training loss for epoch 131: 1.1849294900894165
Training accuracy for epoch 131: 0.6837868690490723
Epoch 132/200
Training loss for epoch 132: 1.1820391416549683

Training accuracy for epoch 132: 0.6842392683029175
Epoch 133/200
Training loss for epoch 133: 1.1791729927062988
Training accuracy for epoch 133: 0.6849840879440308
Epoch 134/200
Training loss for epoch 134: 1.1763306856155396
Training accuracy for epoch 134: 0.6854475736618042
Epoch 135/200
Training loss for epoch 135: 1.1735121011734009
Training accuracy for epoch 135: 0.6858779191970825
Epoch 136/200
Training loss for epoch 136: 1.1707172393798828
Training accuracy for epoch 136: 0.6863413453102112
Epoch 137/200
Training loss for epoch 137: 1.1679449081420898
Training accuracy for epoch 137: 0.6867055296897888
Epoch 138/200
Training loss for epoch 138: 1.1651952266693115
Training accuracy for epoch 138: 0.687185525894165
Epoch 139/200
Training loss for epoch 139: 1.1624679565429688
Training accuracy for epoch 139: 0.6877151727676392
Epoch 140/200
Training loss for epoch 140: 1.1597625017166138
Training accuracy for epoch 140: 0.6881951689720154
Epoch 141/200
Training loss for epoch 141: 1.1570789813995361
Training accuracy for epoch 141: 0.6887413859367371
Epoch 142/200
Training loss for epoch 142: 1.1544164419174194
Training accuracy for epoch 142: 0.6888682842254639
Epoch 143/200
Training loss for epoch 143: 1.1517754793167114
Training accuracy for epoch 143: 0.6893814206123352
Epoch 144/200
Training loss for epoch 144: 1.1491549015045166
Training accuracy for epoch 144: 0.6898945569992065
Epoch 145/200
Training loss for epoch 145: 1.1465550661087036
Training accuracy for epoch 145: 0.6901758909225464
Epoch 146/200
Training loss for epoch 146: 1.1439753770828247
Training accuracy for epoch 146: 0.6906559467315674
Epoch 147/200
Training loss for epoch 147: 1.1414155960083008
Training accuracy for epoch 147: 0.6910366415977478
Epoch 148/200
Training loss for epoch 148: 1.1388753652572632

Training accuracy for epoch 148: 0.6915000677108765
Epoch 149/200
Training loss for epoch 149: 1.1363545656204224
Training accuracy for epoch 149: 0.6918807625770569
Epoch 150/200
Training loss for epoch 150: 1.1338531970977783
Training accuracy for epoch 150: 0.6923276782035828
Epoch 151/200
Training loss for epoch 151: 1.1313704252243042
Training accuracy for epoch 151: 0.6929235458374023
Epoch 152/200
Training loss for epoch 152: 1.12890625
Training accuracy for epoch 152: 0.6934862732887268
Epoch 153/200
Training loss for epoch 153: 1.1264606714248657
Training accuracy for epoch 153: 0.6938835382461548
Epoch 154/200
Training loss for epoch 154: 1.1240336894989014
Training accuracy for epoch 154: 0.6944848895072937
Epoch 155/200
Training loss for epoch 155: 1.1216238737106323
Training accuracy for epoch 155: 0.6948655843734741
Epoch 156/200
Training loss for epoch 156: 1.1192322969436646
Training accuracy for epoch 156: 0.6953455805778503
Epoch 157/200
Training loss for epoch 157: 1.116857886314392
Training accuracy for epoch 157: 0.6955938935279846
Epoch 158/200
Training loss for epoch 158: 1.1145007610321045
Training accuracy for epoch 158: 0.6960076689720154
Epoch 159/200
Training loss for epoch 159: 1.1121609210968018
Training accuracy for epoch 159: 0.6962890625
Epoch 160/200
Training loss for epoch 160: 1.1098377704620361
Training accuracy for epoch 160: 0.6966201066970825
Epoch 161/200
Training loss for epoch 161: 1.1075315475463867
Training accuracy for epoch 161: 0.6970338821411133
Epoch 162/200
Training loss for epoch 162: 1.1052414178848267
Training accuracy for epoch 162: 0.6976132392883301
Epoch 163/200
Training loss for epoch 163: 1.1029677391052246
Training accuracy for epoch 163: 0.6980104446411133
Epoch 164/200
Training loss for epoch 164: 1.1007100343704224

Training accuracy for epoch 164: 0.6985235810279846
Epoch 165/200
Training loss for epoch 165: 1.0984681844711304
Training accuracy for epoch 165: 0.6989207863807678
Epoch 166/200
Training loss for epoch 166: 1.096242070198059
Training accuracy for epoch 166: 0.6993842720985413
Epoch 167/200
Training loss for epoch 167: 1.0940314531326294
Training accuracy for epoch 167: 0.6999304890632629
Epoch 168/200
Training loss for epoch 168: 1.0918360948562622
Training accuracy for epoch 168: 0.7003111839294434
Epoch 169/200
Training loss for epoch 169: 1.0896557569503784
Training accuracy for epoch 169: 0.7007249593734741
Epoch 170/200
Training loss for epoch 170: 1.0874906778335571
Training accuracy for epoch 170: 0.7009401321411133
Epoch 171/200
Training loss for epoch 171: 1.0853400230407715
Training accuracy for epoch 171: 0.7015856504440308
Epoch 172/200
Training loss for epoch 172: 1.0832042694091797
Training accuracy for epoch 172: 0.7019166946411133
Epoch 173/200
Training loss for epoch 173: 1.081082820892334
Training accuracy for epoch 173: 0.7022808790206909
Epoch 174/200
Training loss for epoch 174: 1.0789756774902344
Training accuracy for epoch 174: 0.7027443051338196
Epoch 175/200
Training loss for epoch 175: 1.0768827199935913
Training accuracy for epoch 175: 0.7029925584793091
Epoch 176/200
Training loss for epoch 176: 1.0748037099838257
Training accuracy for epoch 176: 0.7034726142883301
Epoch 177/200
Training loss for epoch 177: 1.0727386474609375
Training accuracy for epoch 177: 0.7040243148803711
Epoch 178/200
Training loss for epoch 178: 1.070687174797058
Training accuracy for epoch 178: 0.7045043110847473
Epoch 179/200
Training loss for epoch 179: 1.068649172782898
Training accuracy for epoch 179: 0.704967737197876
Epoch 180/200
Training loss for epoch 180: 1.0666245222091675

Training accuracy for epoch 180: 0.7053484320640564
Epoch 181/200
Training loss for epoch 181: 1.0646132230758667
Training accuracy for epoch 181: 0.7057291269302368
Epoch 182/200
Training loss for epoch 182: 1.062614917755127
Training accuracy for epoch 182: 0.7062588334083557
Epoch 183/200
Training loss for epoch 183: 1.0606296062469482
Training accuracy for epoch 183: 0.7066560387611389
Epoch 184/200
Training loss for epoch 184: 1.0586570501327515
Training accuracy for epoch 184: 0.7068877816200256
Epoch 185/200
Training loss for epoch 185: 1.0566972494125366
Training accuracy for epoch 185: 0.7070863842964172
Epoch 186/200
Training loss for epoch 186: 1.054749846458435
Training accuracy for epoch 186: 0.7073677778244019
Epoch 187/200
Training loss for epoch 187: 1.0528150796890259
Training accuracy for epoch 187: 0.7077484726905823
Epoch 188/200
Training loss for epoch 188: 1.0508924722671509
Training accuracy for epoch 188: 0.7080132961273193
Epoch 189/200
Training loss for epoch 189: 1.0489819049835205
Training accuracy for epoch 189: 0.7083112597465515
Epoch 190/200
Training loss for epoch 190: 1.0470834970474243
Training accuracy for epoch 190: 0.7086091637611389
Epoch 191/200
Training loss for epoch 191: 1.0451968908309937
Training accuracy for epoch 191: 0.7090560793876648
Epoch 192/200
Training loss for epoch 192: 1.043322205543518
Training accuracy for epoch 192: 0.7093871235847473
Epoch 193/200
Training loss for epoch 193: 1.0414592027664185
Training accuracy for epoch 193: 0.7096353769302368
Epoch 194/200
Training loss for epoch 194: 1.0396075248718262
Training accuracy for epoch 194: 0.7099829912185669
Epoch 195/200
Training loss for epoch 195: 1.0377674102783203
Training accuracy for epoch 195: 0.710198163986206
Epoch 196/200
Training loss for epoch 196: 1.0359386205673218

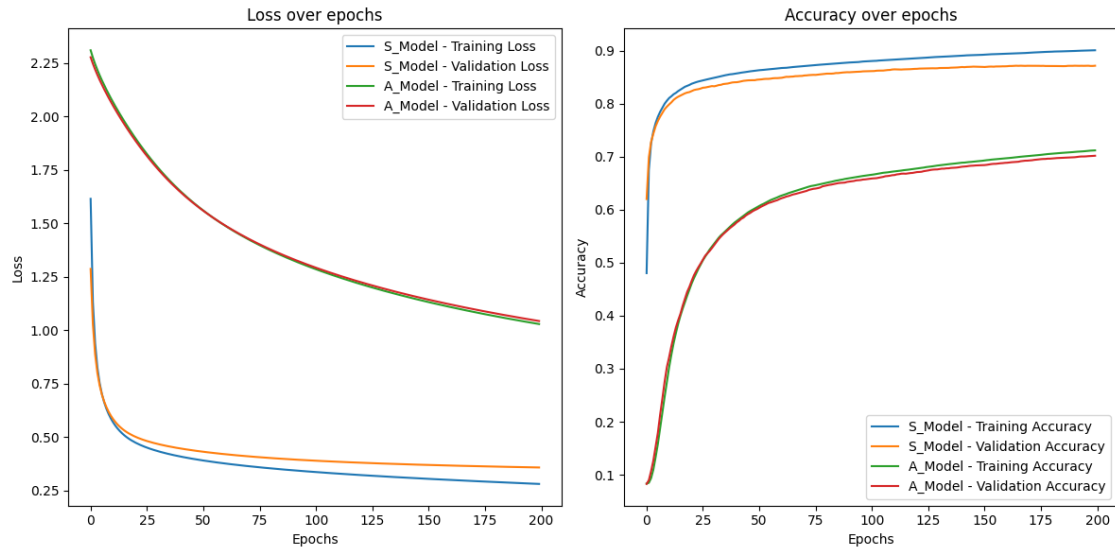
Training accuracy for epoch 196: 0.7106615900993347
Epoch 197/200
Training loss for epoch 197: 1.0341211557388306
Training accuracy for epoch 197: 0.7110092043876648
Epoch 198/200
Training loss for epoch 198: 1.0323145389556885
Training accuracy for epoch 198: 0.7114561200141907
Epoch 199/200
Training loss for epoch 199: 1.0305190086364746
Training accuracy for epoch 199: 0.7116547226905823
Epoch 200/200
Training loss for epoch 200: 1.0287344455718994
Training accuracy for epoch 200: 0.711886465549469

```
[ ]: # Plot training and validation loss over each epoch
plt.figure(figsize=(12, 6))

plt.subplot(1, 2, 1)
plt.plot(train_losses_standard_fashion, label='S_Model - Training Loss')
plt.plot(val_losses_standard_fashion, label='S_Model - Validation Loss')
plt.plot(a_train_losses_fashion, label='A_Model - Training Loss')
plt.plot(a_val_losses_fashion, label='A_Model - Validation Loss')
plt.title('Loss over epochs')
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.legend()

plt.subplot(1, 2, 2)
plt.plot(train_accuracies_standard_fashion, label='S_Model - Training Accuracy')
plt.plot(val_accuracies_standard_fashion, label='S_Model - Validation Accuracy')
plt.plot(a_train_accuracies_fashion, label='A_Model - Training Accuracy')
plt.plot(a_val_accuracies_fashion, label='A_Model - Validation Accuracy')
plt.title('Accuracy over epochs')
plt.xlabel('Epochs')
plt.ylabel('Accuracy')
plt.legend()

plt.tight_layout()
plt.show()
```



```
[ ]: from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
import seaborn as sns

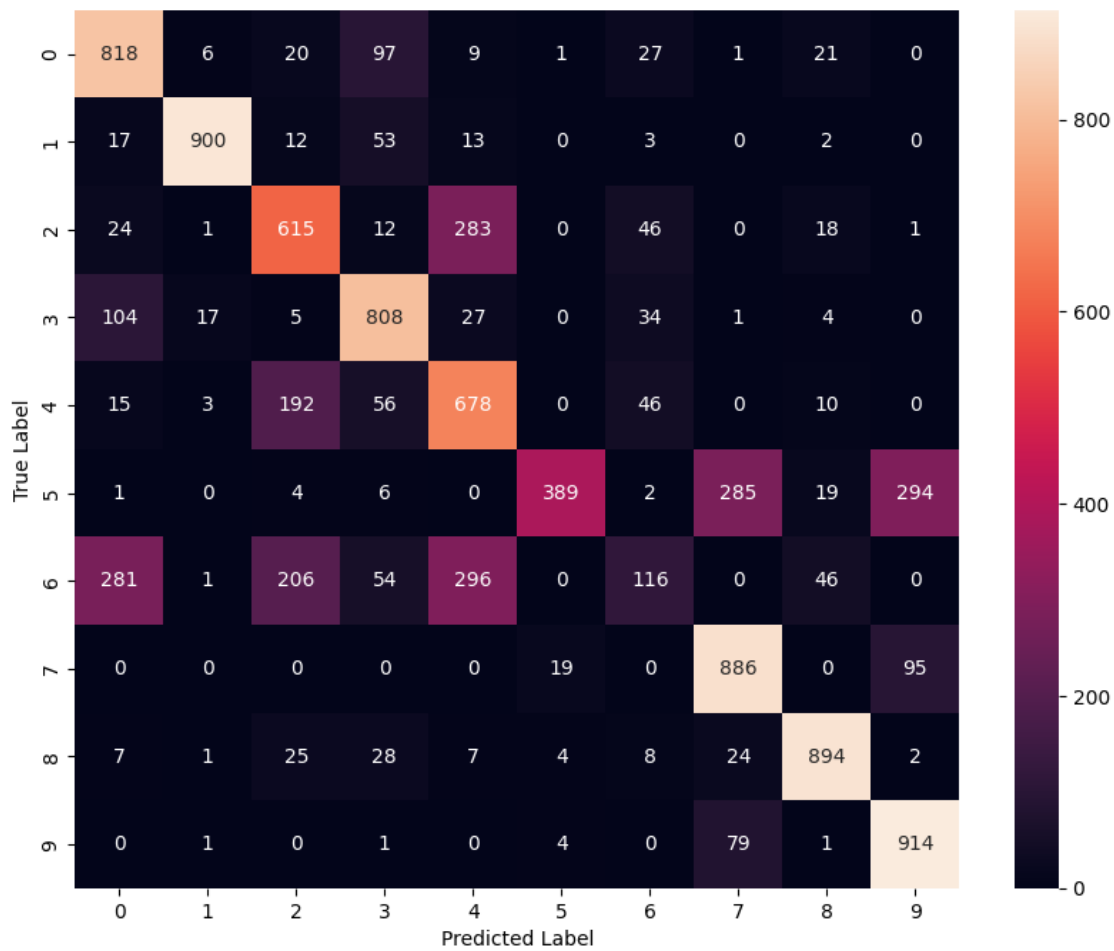
# Predict the values from the validation dataset
y_pred = a2_model.predict(X_test_fashion)
# Convert predictions classes to one hot vectors
y_pred_classes = np.argmax(y_pred, axis = 1)

# compute the confusion matrix
confusion_mtx = confusion_matrix(y_test_fashion, y_pred_classes)

# plot the confusion matrix
plt.figure(figsize=(10,8))
sns.heatmap(confusion_mtx, annot=True, fmt='d')
plt.ylabel('True Label')
plt.xlabel('Predicted Label')
plt.show()

# Classification report
print(classification_report(y_test_fashion, y_pred_classes))
```

313/313 [=====] - 1s 2ms/step



	precision	recall	f1-score	support
0	0.65	0.82	0.72	1000
1	0.97	0.90	0.93	1000
2	0.57	0.61	0.59	1000
3	0.72	0.81	0.76	1000
4	0.52	0.68	0.59	1000
5	0.93	0.39	0.55	1000
6	0.41	0.12	0.18	1000
7	0.69	0.89	0.78	1000
8	0.88	0.89	0.89	1000
9	0.70	0.91	0.79	1000
accuracy			0.70	10000
macro avg	0.70	0.70	0.68	10000
weighted avg	0.70	0.70	0.68	10000

10 Discussion of Results

The visualizations of the loss and accuracy curves show us that both training approaches across both data sets result in different “learning styles”. We can see that the standard model seems to learn very quickly, and performs very well on both the training and validation data. In comparison, the accumulation model learns at a slower, yet very consistent rate, and also shows a similar performance on both the training and validation data. This is as expected because the standard model simply has more optimization steps (every batch) compared to one optimization step every epoch with the accumulation model. The smooth curves tell us that the training process is stable and since there is basically no divergence between training and validation data, the model is far from overfitting. Overall, both models are training normally, and the graphs look as expected. Obviously it is worth pointing out that the standard model performs better overall (lower cost and higher accuracy). However, the fact that the gradient accumulation approach also gives us a good result is promising.