

In [1]:

```
import tensorflow as tf
#import tensorflow_datasets as tfds
```

Baseline model

In [2]:

```
# Load Data
# tfds.list_builders()
# (train, test), info = tfds.load("mnist", split=['train', 'test'], with_info=True, as_supervised=True)
(x_train, y_train), (x_test, y_test) = tf.keras.datasets.fashion_mnist.load_data()
```

```
Downloading data from https://storage.googleapis.com/tensorflow/tf-keras-datasets/train-labels-idx1-
ubyte.gz
32768/29515 [=====] - 0s 0us/step
40960/29515 [=====] - 0s 0us/step
Downloading data from https://storage.googleapis.com/tensorflow/tf-keras-datasets/train-images-idx3-
ubyte.gz
26427392/26421880 [=====] - 0s 0us/step
26435584/26421880 [=====] - 0s 0us/step
Downloading data from https://storage.googleapis.com/tensorflow/tf-keras-datasets/t10k-labels-idx1-u
byte.gz
16384/5148 [=====] - 0s 0us/step
Downloading data from https://storage.googleapis.com/tensorflow/tf-keras-datasets/t10k-images-idx3-u
byte.gz
4423680/4422102 [=====] - 0s 0us/step
4431872/4422102 [=====] - 0s 0us/step
```

In [3]:

```
# Baseline model definition
model = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(16, activation='relu'),
    tf.keras.layers.Dense(16, activation='relu'),
    tf.keras.layers.Dense(10, activation='softmax')
])
```

In [4]:

```
# Baseline model compilation
model.compile(optimizer='adam',
              loss='sparse_categorical_crossentropy',
              metrics=['sparse_categorical_accuracy'])
```

In [5]:

```
# Baseline model fitting
history = model.fit(x_train, y_train,
                    batch_size=128,
                    epochs=100,
                    validation_data=(x_test, y_test),
                    verbose=2
                    )
```

```
Epoch 1/100
469/469 - 1s - loss: 2.4366 - sparse_categorical_accuracy: 0.3531 - val_loss: 1.5686 - val_sparse_ca
tegorical_accuracy: 0.4122 - 1s/epoch - 3ms/step
Epoch 2/100
469/469 - 1s - loss: 1.3791 - sparse_categorical_accuracy: 0.4474 - val_loss: 1.3388 - val_sparse_ca
tegorical_accuracy: 0.5069 - 819ms/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 1.1816 - sparse_categorical_accuracy: 0.5167 - val_loss: 1.1218 - val_sparse_ca
tegorical_accuracy: 0.5218 - 717ms/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 0.9765 - sparse_categorical_accuracy: 0.6170 - val_loss: 0.8656 - val_sparse_ca
tegorical_accuracy: 0.6768 - 806ms/epoch - 2ms/step
Epoch 5/100
469/469 - 2s - loss: 0.7980 - sparse_categorical_accuracy: 0.6909 - val_loss: 0.7821 - val_sparse_ca
tegorical_accuracy: 0.6860 - 2s/epoch - 3ms/step
Epoch 6/100
469/469 - 2s - loss: 0.7271 - sparse_categorical_accuracy: 0.7056 - val_loss: 0.7179 - val_sparse_ca
tegorical_accuracy: 0.7076 - 2s/epoch - 3ms/step
Epoch 7/100
469/469 - 1s - loss: 0.6944 - sparse_categorical_accuracy: 0.7169 - val_loss: 0.6976 - val_sparse_ca
tegorical_accuracy: 0.7156 - 1s/epoch - 3ms/step
```

Epoch 8/100
469/469 - 1s - loss: 0.6572 - sparse_categorical_accuracy: 0.7359 - val_loss: 0.6684 - val_sparse_categorical_accuracy: 0.7483 - 1s/epoch - 3ms/step

Epoch 9/100
469/469 - 1s - loss: 0.6204 - sparse_categorical_accuracy: 0.7621 - val_loss: 0.6665 - val_sparse_categorical_accuracy: 0.7564 - 1s/epoch - 3ms/step

Epoch 10/100
469/469 - 1s - loss: 0.6069 - sparse_categorical_accuracy: 0.7703 - val_loss: 0.6287 - val_sparse_categorical_accuracy: 0.7637 - 1s/epoch - 3ms/step

Epoch 11/100
469/469 - 1s - loss: 0.5922 - sparse_categorical_accuracy: 0.7773 - val_loss: 0.6453 - val_sparse_categorical_accuracy: 0.7609 - 780ms/epoch - 2ms/step

Epoch 12/100
469/469 - 1s - loss: 0.5746 - sparse_categorical_accuracy: 0.7806 - val_loss: 0.6209 - val_sparse_categorical_accuracy: 0.7643 - 798ms/epoch - 2ms/step

Epoch 13/100
469/469 - 1s - loss: 0.5649 - sparse_categorical_accuracy: 0.7829 - val_loss: 0.6428 - val_sparse_categorical_accuracy: 0.7695 - 707ms/epoch - 2ms/step

Epoch 14/100
469/469 - 1s - loss: 0.5624 - sparse_categorical_accuracy: 0.7871 - val_loss: 0.6087 - val_sparse_categorical_accuracy: 0.7726 - 727ms/epoch - 2ms/step

Epoch 15/100
469/469 - 1s - loss: 0.5564 - sparse_categorical_accuracy: 0.7872 - val_loss: 0.5871 - val_sparse_categorical_accuracy: 0.7834 - 735ms/epoch - 2ms/step

Epoch 16/100
469/469 - 1s - loss: 0.5525 - sparse_categorical_accuracy: 0.7879 - val_loss: 0.6291 - val_sparse_categorical_accuracy: 0.7636 - 767ms/epoch - 2ms/step

Epoch 17/100
469/469 - 1s - loss: 0.5421 - sparse_categorical_accuracy: 0.7900 - val_loss: 0.6231 - val_sparse_categorical_accuracy: 0.7742 - 756ms/epoch - 2ms/step

Epoch 18/100
469/469 - 1s - loss: 0.5379 - sparse_categorical_accuracy: 0.7917 - val_loss: 0.5779 - val_sparse_categorical_accuracy: 0.7826 - 731ms/epoch - 2ms/step

Epoch 19/100
469/469 - 1s - loss: 0.5382 - sparse_categorical_accuracy: 0.7928 - val_loss: 0.5891 - val_sparse_categorical_accuracy: 0.7841 - 730ms/epoch - 2ms/step

Epoch 20/100
469/469 - 1s - loss: 0.5304 - sparse_categorical_accuracy: 0.7952 - val_loss: 0.5812 - val_sparse_categorical_accuracy: 0.7834 - 803ms/epoch - 2ms/step

Epoch 21/100
469/469 - 1s - loss: 0.5279 - sparse_categorical_accuracy: 0.7957 - val_loss: 0.5994 - val_sparse_categorical_accuracy: 0.7761 - 809ms/epoch - 2ms/step

Epoch 22/100
469/469 - 1s - loss: 0.5222 - sparse_categorical_accuracy: 0.7966 - val_loss: 0.5816 - val_sparse_categorical_accuracy: 0.7825 - 718ms/epoch - 2ms/step

Epoch 23/100
469/469 - 1s - loss: 0.5195 - sparse_categorical_accuracy: 0.7983 - val_loss: 0.5843 - val_sparse_categorical_accuracy: 0.7837 - 790ms/epoch - 2ms/step

Epoch 24/100
469/469 - 1s - loss: 0.5152 - sparse_categorical_accuracy: 0.8000 - val_loss: 0.5844 - val_sparse_categorical_accuracy: 0.7857 - 734ms/epoch - 2ms/step

Epoch 25/100
469/469 - 1s - loss: 0.5203 - sparse_categorical_accuracy: 0.8003 - val_loss: 0.5599 - val_sparse_categorical_accuracy: 0.7861 - 794ms/epoch - 2ms/step

Epoch 26/100
469/469 - 1s - loss: 0.5144 - sparse_categorical_accuracy: 0.8009 - val_loss: 0.5974 - val_sparse_categorical_accuracy: 0.7801 - 712ms/epoch - 2ms/step

Epoch 27/100
469/469 - 1s - loss: 0.5159 - sparse_categorical_accuracy: 0.8012 - val_loss: 0.5814 - val_sparse_categorical_accuracy: 0.7814 - 718ms/epoch - 2ms/step

Epoch 28/100
469/469 - 1s - loss: 0.5086 - sparse_categorical_accuracy: 0.8008 - val_loss: 0.5579 - val_sparse_categorical_accuracy: 0.7904 - 779ms/epoch - 2ms/step

Epoch 29/100
469/469 - 1s - loss: 0.5068 - sparse_categorical_accuracy: 0.8037 - val_loss: 0.5862 - val_sparse_categorical_accuracy: 0.7821 - 793ms/epoch - 2ms/step

Epoch 30/100
469/469 - 1s - loss: 0.5031 - sparse_categorical_accuracy: 0.8045 - val_loss: 0.5613 - val_sparse_categorical_accuracy: 0.7865 - 784ms/epoch - 2ms/step

Epoch 31/100
469/469 - 1s - loss: 0.5066 - sparse_categorical_accuracy: 0.8031 - val_loss: 0.5690 - val_sparse_categorical_accuracy: 0.7828 - 750ms/epoch - 2ms/step

Epoch 32/100
469/469 - 1s - loss: 0.5063 - sparse_categorical_accuracy: 0.8030 - val_loss: 0.5665 - val_sparse_categorical_accuracy: 0.7905 - 707ms/epoch - 2ms/step

Epoch 33/100
469/469 - 1s - loss: 0.5061 - sparse_categorical_accuracy: 0.8032 - val_loss: 0.5580 - val_sparse_categorical_accuracy: 0.7881 - 738ms/epoch - 2ms/step

Epoch 34/100
469/469 - 1s - loss: 0.4987 - sparse_categorical_accuracy: 0.8062 - val_loss: 0.5737 - val_sparse_categorical_accuracy: 0.7803 - 794ms/epoch - 2ms/step

Epoch 35/100
469/469 - 1s - loss: 0.4973 - sparse_categorical_accuracy: 0.8061 - val_loss: 0.5503 - val_sparse_categorical_accuracy: 0.7803 - 794ms/epoch - 2ms/step

tegorical_accuracy: 0.7916 - 713ms/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.4884 - sparse_categorical_accuracy: 0.8085 - val_loss: 0.5500 - val_sparse_ca
tegorical_accuracy: 0.7939 - 727ms/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.4972 - sparse_categorical_accuracy: 0.8059 - val_loss: 0.5708 - val_sparse_ca
tegorical_accuracy: 0.7879 - 712ms/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.4972 - sparse_categorical_accuracy: 0.8077 - val_loss: 0.8575 - val_sparse_ca
tegorical_accuracy: 0.7470 - 766ms/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.4976 - sparse_categorical_accuracy: 0.8065 - val_loss: 0.5886 - val_sparse_ca
tegorical_accuracy: 0.7803 - 835ms/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 0.4878 - sparse_categorical_accuracy: 0.8082 - val_loss: 0.5658 - val_sparse_ca
tegorical_accuracy: 0.7892 - 775ms/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.4869 - sparse_categorical_accuracy: 0.8091 - val_loss: 0.5763 - val_sparse_ca
tegorical_accuracy: 0.7822 - 796ms/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.4922 - sparse_categorical_accuracy: 0.8070 - val_loss: 0.5459 - val_sparse_ca
tegorical_accuracy: 0.7947 - 700ms/epoch - 1ms/step
Epoch 43/100
469/469 - 1s - loss: 0.5091 - sparse_categorical_accuracy: 0.8052 - val_loss: 0.5867 - val_sparse_ca
tegorical_accuracy: 0.7838 - 711ms/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.4961 - sparse_categorical_accuracy: 0.8059 - val_loss: 0.5470 - val_sparse_ca
tegorical_accuracy: 0.7927 - 744ms/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.4808 - sparse_categorical_accuracy: 0.8088 - val_loss: 0.5427 - val_sparse_ca
tegorical_accuracy: 0.7969 - 730ms/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.4917 - sparse_categorical_accuracy: 0.8083 - val_loss: 0.5440 - val_sparse_ca
tegorical_accuracy: 0.7921 - 773ms/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.4852 - sparse_categorical_accuracy: 0.8089 - val_loss: 0.5467 - val_sparse_ca
tegorical_accuracy: 0.7903 - 764ms/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.4922 - sparse_categorical_accuracy: 0.8073 - val_loss: 0.5541 - val_sparse_ca
tegorical_accuracy: 0.7869 - 774ms/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.4851 - sparse_categorical_accuracy: 0.8090 - val_loss: 0.5670 - val_sparse_ca
tegorical_accuracy: 0.7889 - 759ms/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.4817 - sparse_categorical_accuracy: 0.8094 - val_loss: 0.5632 - val_sparse_ca
tegorical_accuracy: 0.7894 - 824ms/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.4735 - sparse_categorical_accuracy: 0.8115 - val_loss: 0.5485 - val_sparse_ca
tegorical_accuracy: 0.7928 - 777ms/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.4774 - sparse_categorical_accuracy: 0.8115 - val_loss: 0.5730 - val_sparse_ca
tegorical_accuracy: 0.7843 - 796ms/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.4890 - sparse_categorical_accuracy: 0.8080 - val_loss: 0.5645 - val_sparse_ca
tegorical_accuracy: 0.7882 - 712ms/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.4882 - sparse_categorical_accuracy: 0.8064 - val_loss: 0.5525 - val_sparse_ca
tegorical_accuracy: 0.7901 - 717ms/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.4799 - sparse_categorical_accuracy: 0.8100 - val_loss: 0.5772 - val_sparse_ca
tegorical_accuracy: 0.7837 - 721ms/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.4720 - sparse_categorical_accuracy: 0.8138 - val_loss: 0.5433 - val_sparse_ca
tegorical_accuracy: 0.7951 - 727ms/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 0.4746 - sparse_categorical_accuracy: 0.8123 - val_loss: 0.5485 - val_sparse_ca
tegorical_accuracy: 0.7935 - 752ms/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.4759 - sparse_categorical_accuracy: 0.8104 - val_loss: 0.5639 - val_sparse_ca
tegorical_accuracy: 0.7902 - 711ms/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.4814 - sparse_categorical_accuracy: 0.8104 - val_loss: 0.5768 - val_sparse_ca
tegorical_accuracy: 0.7851 - 789ms/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.4730 - sparse_categorical_accuracy: 0.8114 - val_loss: 0.5620 - val_sparse_ca
tegorical_accuracy: 0.7865 - 795ms/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.4800 - sparse_categorical_accuracy: 0.8099 - val_loss: 0.5444 - val_sparse_ca
tegorical_accuracy: 0.7933 - 745ms/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.4698 - sparse_categorical_accuracy: 0.8122 - val_loss: 0.5638 - val_sparse_ca
tegorical_accuracy: 0.7871 - 786ms/epoch - 2ms/step
Epoch 63/100

469/469 - 1s - loss: 0.4779 - sparse_categorical_accuracy: 0.8119 - val_loss: 0.5511 - val_sparse_categorical_accuracy: 0.7879 - 719ms/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.4709 - sparse_categorical_accuracy: 0.8122 - val_loss: 0.5445 - val_sparse_categorical_accuracy: 0.7943 - 796ms/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.4706 - sparse_categorical_accuracy: 0.8139 - val_loss: 0.5371 - val_sparse_categorical_accuracy: 0.7957 - 746ms/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.4739 - sparse_categorical_accuracy: 0.8112 - val_loss: 0.5414 - val_sparse_categorical_accuracy: 0.7903 - 717ms/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.4655 - sparse_categorical_accuracy: 0.8141 - val_loss: 0.5444 - val_sparse_categorical_accuracy: 0.7921 - 800ms/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 0.4707 - sparse_categorical_accuracy: 0.8118 - val_loss: 0.5506 - val_sparse_categorical_accuracy: 0.7958 - 800ms/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 0.4636 - sparse_categorical_accuracy: 0.8145 - val_loss: 0.5372 - val_sparse_categorical_accuracy: 0.7969 - 726ms/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.4709 - sparse_categorical_accuracy: 0.8133 - val_loss: 0.5487 - val_sparse_categorical_accuracy: 0.7969 - 752ms/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.4724 - sparse_categorical_accuracy: 0.8102 - val_loss: 0.5654 - val_sparse_categorical_accuracy: 0.7848 - 821ms/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.4729 - sparse_categorical_accuracy: 0.8117 - val_loss: 0.5509 - val_sparse_categorical_accuracy: 0.7909 - 710ms/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.4597 - sparse_categorical_accuracy: 0.8154 - val_loss: 0.5319 - val_sparse_categorical_accuracy: 0.7979 - 711ms/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.4695 - sparse_categorical_accuracy: 0.8115 - val_loss: 0.5233 - val_sparse_categorical_accuracy: 0.7966 - 729ms/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.4608 - sparse_categorical_accuracy: 0.8141 - val_loss: 0.5549 - val_sparse_categorical_accuracy: 0.7898 - 762ms/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.4663 - sparse_categorical_accuracy: 0.8128 - val_loss: 0.5513 - val_sparse_categorical_accuracy: 0.7900 - 796ms/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.4728 - sparse_categorical_accuracy: 0.8118 - val_loss: 0.5326 - val_sparse_categorical_accuracy: 0.7962 - 776ms/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.4590 - sparse_categorical_accuracy: 0.8156 - val_loss: 0.5264 - val_sparse_categorical_accuracy: 0.7968 - 754ms/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.4635 - sparse_categorical_accuracy: 0.8140 - val_loss: 0.5402 - val_sparse_categorical_accuracy: 0.7959 - 779ms/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.4593 - sparse_categorical_accuracy: 0.8147 - val_loss: 0.5569 - val_sparse_categorical_accuracy: 0.7876 - 714ms/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.4616 - sparse_categorical_accuracy: 0.8136 - val_loss: 0.5353 - val_sparse_categorical_accuracy: 0.7957 - 704ms/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.4640 - sparse_categorical_accuracy: 0.8138 - val_loss: 0.5402 - val_sparse_categorical_accuracy: 0.7942 - 723ms/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.4584 - sparse_categorical_accuracy: 0.8151 - val_loss: 0.5345 - val_sparse_categorical_accuracy: 0.7977 - 797ms/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.4579 - sparse_categorical_accuracy: 0.8155 - val_loss: 0.5603 - val_sparse_categorical_accuracy: 0.7959 - 788ms/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.4608 - sparse_categorical_accuracy: 0.8175 - val_loss: 0.5351 - val_sparse_categorical_accuracy: 0.7914 - 801ms/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.4579 - sparse_categorical_accuracy: 0.8145 - val_loss: 0.5407 - val_sparse_categorical_accuracy: 0.7928 - 718ms/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.4665 - sparse_categorical_accuracy: 0.8138 - val_loss: 0.5451 - val_sparse_categorical_accuracy: 0.7967 - 767ms/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.4620 - sparse_categorical_accuracy: 0.8152 - val_loss: 0.6235 - val_sparse_categorical_accuracy: 0.7756 - 809ms/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.4605 - sparse_categorical_accuracy: 0.8166 - val_loss: 0.5363 - val_sparse_categorical_accuracy: 0.7961 - 752ms/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.4576 - sparse_categorical_accuracy: 0.8159 - val_loss: 0.5518 - val_sparse_categorical_accuracy: 0.7789 - 797ms/epoch - 2ms/step

```
Epoch 91/100
469/469 - 1s - loss: 0.4630 - sparse_categorical_accuracy: 0.8130 - val_loss: 0.5670 - val_sparse_categorical_accuracy: 0.7847 - 809ms/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.4641 - sparse_categorical_accuracy: 0.8138 - val_loss: 0.5333 - val_sparse_categorical_accuracy: 0.7976 - 789ms/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.4645 - sparse_categorical_accuracy: 0.8160 - val_loss: 0.5380 - val_sparse_categorical_accuracy: 0.7966 - 733ms/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.4616 - sparse_categorical_accuracy: 0.8154 - val_loss: 0.5412 - val_sparse_categorical_accuracy: 0.7941 - 740ms/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.4599 - sparse_categorical_accuracy: 0.8183 - val_loss: 0.5366 - val_sparse_categorical_accuracy: 0.7948 - 743ms/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.4587 - sparse_categorical_accuracy: 0.8185 - val_loss: 0.5401 - val_sparse_categorical_accuracy: 0.7939 - 936ms/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.4584 - sparse_categorical_accuracy: 0.8161 - val_loss: 0.5482 - val_sparse_categorical_accuracy: 0.7952 - 931ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.4646 - sparse_categorical_accuracy: 0.8159 - val_loss: 0.5413 - val_sparse_categorical_accuracy: 0.7901 - 917ms/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.4692 - sparse_categorical_accuracy: 0.8145 - val_loss: 0.5818 - val_sparse_categorical_accuracy: 0.7802 - 713ms/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.4649 - sparse_categorical_accuracy: 0.8136 - val_loss: 0.5428 - val_sparse_categorical_accuracy: 0.7941 - 778ms/epoch - 2ms/step
```

In [6]:

```
# Baseline model evaluation
model.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.5428 - sparse_categorical_accuracy: 0.7941 - 226ms/epoch - 723us/step
```

Out[6]:

```
[0.5427749752998352, 0.7940999865531921]
```

The baseline model with `kernel_initializer='glorot_uniform'`, `bias_initializer='zeros'`

Try different kernel initializers

Model with `glorot_normal` kernel initializer

In [7]:

```
model_kernel1 = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(16, activation='relu', kernel_initializer='glorot_normal'),
    tf.keras.layers.Dense(16, activation='relu', kernel_initializer='glorot_normal'),
    tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='glorot_normal')
])
```

In [8]:

```
model_kernel1.compile(optimizer='adam',
                      loss='sparse_categorical_crossentropy',
                      metrics=['sparse_categorical_accuracy'])
```

In [9]:

```
history_kernel1 = model_kernel1.fit(x_train, y_train,
                                    batch_size=128,
                                    epochs=100,
                                    validation_data=(x_test, y_test),
                                    verbose=2)
```

```
Epoch 1/100
469/469 - 1s - loss: 3.5399 - sparse_categorical_accuracy: 0.3419 - val_loss: 1.5738 - val_sparse_categorical_accuracy: 0.4320 - 1s/epoch - 2ms/step
Epoch 2/100
469/469 - 1s - loss: 1.4560 - sparse_categorical_accuracy: 0.4459 - val_loss: 1.3433 - val_sparse_categorical_accuracy: 0.4633 - 804ms/epoch - 2ms/step
Epoch 3/100
```

469/469 - 1s - loss: 1.3017 - sparse_categorical_accuracy: 0.4783 - val_loss: 1.2152 - val_sparse_categorical_accuracy: 0.5373 - 802ms/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 1.1285 - sparse_categorical_accuracy: 0.5589 - val_loss: 1.0768 - val_sparse_categorical_accuracy: 0.5763 - 709ms/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 1.0563 - sparse_categorical_accuracy: 0.5775 - val_loss: 1.0580 - val_sparse_categorical_accuracy: 0.5709 - 809ms/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 1.0069 - sparse_categorical_accuracy: 0.5969 - val_loss: 1.0071 - val_sparse_categorical_accuracy: 0.5931 - 724ms/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.9346 - sparse_categorical_accuracy: 0.6372 - val_loss: 0.8998 - val_sparse_categorical_accuracy: 0.6656 - 796ms/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 0.8162 - sparse_categorical_accuracy: 0.7032 - val_loss: 0.8054 - val_sparse_categorical_accuracy: 0.7123 - 767ms/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 0.7384 - sparse_categorical_accuracy: 0.7323 - val_loss: 0.7436 - val_sparse_categorical_accuracy: 0.7272 - 728ms/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 0.6710 - sparse_categorical_accuracy: 0.7577 - val_loss: 0.6638 - val_sparse_categorical_accuracy: 0.7660 - 720ms/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 0.6011 - sparse_categorical_accuracy: 0.7848 - val_loss: 0.6238 - val_sparse_categorical_accuracy: 0.7854 - 794ms/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.5705 - sparse_categorical_accuracy: 0.7961 - val_loss: 0.6031 - val_sparse_categorical_accuracy: 0.7924 - 779ms/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.5488 - sparse_categorical_accuracy: 0.8032 - val_loss: 0.5908 - val_sparse_categorical_accuracy: 0.7951 - 800ms/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.5322 - sparse_categorical_accuracy: 0.8111 - val_loss: 0.5752 - val_sparse_categorical_accuracy: 0.8017 - 703ms/epoch - 1ms/step
Epoch 15/100
469/469 - 1s - loss: 0.5222 - sparse_categorical_accuracy: 0.8146 - val_loss: 0.5922 - val_sparse_categorical_accuracy: 0.7894 - 780ms/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.5169 - sparse_categorical_accuracy: 0.8153 - val_loss: 0.5889 - val_sparse_categorical_accuracy: 0.8022 - 797ms/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.5037 - sparse_categorical_accuracy: 0.8192 - val_loss: 0.5500 - val_sparse_categorical_accuracy: 0.8126 - 723ms/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 0.4947 - sparse_categorical_accuracy: 0.8236 - val_loss: 0.5513 - val_sparse_categorical_accuracy: 0.8135 - 804ms/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.4884 - sparse_categorical_accuracy: 0.8271 - val_loss: 0.5379 - val_sparse_categorical_accuracy: 0.8188 - 714ms/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.4788 - sparse_categorical_accuracy: 0.8308 - val_loss: 0.5329 - val_sparse_categorical_accuracy: 0.8193 - 784ms/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.4740 - sparse_categorical_accuracy: 0.8311 - val_loss: 0.5192 - val_sparse_categorical_accuracy: 0.8236 - 720ms/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.4680 - sparse_categorical_accuracy: 0.8324 - val_loss: 0.5216 - val_sparse_categorical_accuracy: 0.8232 - 749ms/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.4598 - sparse_categorical_accuracy: 0.8357 - val_loss: 0.5167 - val_sparse_categorical_accuracy: 0.8194 - 721ms/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.4581 - sparse_categorical_accuracy: 0.8372 - val_loss: 0.5167 - val_sparse_categorical_accuracy: 0.8256 - 717ms/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.4422 - sparse_categorical_accuracy: 0.8429 - val_loss: 0.5061 - val_sparse_categorical_accuracy: 0.8318 - 748ms/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.4433 - sparse_categorical_accuracy: 0.8439 - val_loss: 0.4959 - val_sparse_categorical_accuracy: 0.8347 - 737ms/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.4346 - sparse_categorical_accuracy: 0.8464 - val_loss: 0.5131 - val_sparse_categorical_accuracy: 0.8266 - 791ms/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.4334 - sparse_categorical_accuracy: 0.8474 - val_loss: 0.5082 - val_sparse_categorical_accuracy: 0.8282 - 704ms/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.4218 - sparse_categorical_accuracy: 0.8529 - val_loss: 0.4992 - val_sparse_categorical_accuracy: 0.8333 - 736ms/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.4210 - sparse_categorical_accuracy: 0.8525 - val_loss: 0.5041 - val_sparse_categorical_accuracy: 0.8359 - 724ms/epoch - 2ms/step

Epoch 31/100
469/469 - 1s - loss: 0.4138 - sparse_categorical_accuracy: 0.8569 - val_loss: 0.4974 - val_sparse_categorical_accuracy: 0.8366 - 732ms/epoch - 2ms/step

Epoch 32/100
469/469 - 1s - loss: 0.4107 - sparse_categorical_accuracy: 0.8564 - val_loss: 0.4858 - val_sparse_categorical_accuracy: 0.8396 - 806ms/epoch - 2ms/step

Epoch 33/100
469/469 - 1s - loss: 0.4083 - sparse_categorical_accuracy: 0.8558 - val_loss: 0.5038 - val_sparse_categorical_accuracy: 0.8289 - 791ms/epoch - 2ms/step

Epoch 34/100
469/469 - 1s - loss: 0.4034 - sparse_categorical_accuracy: 0.8586 - val_loss: 0.4991 - val_sparse_categorical_accuracy: 0.8332 - 794ms/epoch - 2ms/step

Epoch 35/100
469/469 - 1s - loss: 0.3976 - sparse_categorical_accuracy: 0.8611 - val_loss: 0.4954 - val_sparse_categorical_accuracy: 0.8388 - 760ms/epoch - 2ms/step

Epoch 36/100
469/469 - 1s - loss: 0.3926 - sparse_categorical_accuracy: 0.8622 - val_loss: 0.4847 - val_sparse_categorical_accuracy: 0.8406 - 726ms/epoch - 2ms/step

Epoch 37/100
469/469 - 1s - loss: 0.3886 - sparse_categorical_accuracy: 0.8634 - val_loss: 0.4868 - val_sparse_categorical_accuracy: 0.8416 - 780ms/epoch - 2ms/step

Epoch 38/100
469/469 - 1s - loss: 0.3844 - sparse_categorical_accuracy: 0.8656 - val_loss: 0.4688 - val_sparse_categorical_accuracy: 0.8473 - 794ms/epoch - 2ms/step

Epoch 39/100
469/469 - 1s - loss: 0.3869 - sparse_categorical_accuracy: 0.8638 - val_loss: 0.4880 - val_sparse_categorical_accuracy: 0.8393 - 790ms/epoch - 2ms/step

Epoch 40/100
469/469 - 1s - loss: 0.3789 - sparse_categorical_accuracy: 0.8672 - val_loss: 0.4999 - val_sparse_categorical_accuracy: 0.8324 - 692ms/epoch - 1ms/step

Epoch 41/100
469/469 - 1s - loss: 0.3786 - sparse_categorical_accuracy: 0.8673 - val_loss: 0.4968 - val_sparse_categorical_accuracy: 0.8347 - 763ms/epoch - 2ms/step

Epoch 42/100
469/469 - 1s - loss: 0.3760 - sparse_categorical_accuracy: 0.8672 - val_loss: 0.5222 - val_sparse_categorical_accuracy: 0.8292 - 743ms/epoch - 2ms/step

Epoch 43/100
469/469 - 1s - loss: 0.3775 - sparse_categorical_accuracy: 0.8677 - val_loss: 0.4896 - val_sparse_categorical_accuracy: 0.8381 - 828ms/epoch - 2ms/step

Epoch 44/100
469/469 - 1s - loss: 0.3728 - sparse_categorical_accuracy: 0.8682 - val_loss: 0.4692 - val_sparse_categorical_accuracy: 0.8504 - 812ms/epoch - 2ms/step

Epoch 45/100
469/469 - 1s - loss: 0.3681 - sparse_categorical_accuracy: 0.8687 - val_loss: 0.4829 - val_sparse_categorical_accuracy: 0.8473 - 733ms/epoch - 2ms/step

Epoch 46/100
469/469 - 1s - loss: 0.3704 - sparse_categorical_accuracy: 0.8689 - val_loss: 0.4688 - val_sparse_categorical_accuracy: 0.8487 - 841ms/epoch - 2ms/step

Epoch 47/100
469/469 - 1s - loss: 0.3704 - sparse_categorical_accuracy: 0.8694 - val_loss: 0.5104 - val_sparse_categorical_accuracy: 0.8304 - 723ms/epoch - 2ms/step

Epoch 48/100
469/469 - 1s - loss: 0.3638 - sparse_categorical_accuracy: 0.8710 - val_loss: 0.4887 - val_sparse_categorical_accuracy: 0.8410 - 743ms/epoch - 2ms/step

Epoch 49/100
469/469 - 1s - loss: 0.3634 - sparse_categorical_accuracy: 0.8719 - val_loss: 0.4809 - val_sparse_categorical_accuracy: 0.8449 - 700ms/epoch - 1ms/step

Epoch 50/100
469/469 - 1s - loss: 0.3623 - sparse_categorical_accuracy: 0.8715 - val_loss: 0.4873 - val_sparse_categorical_accuracy: 0.8447 - 773ms/epoch - 2ms/step

Epoch 51/100
469/469 - 1s - loss: 0.3632 - sparse_categorical_accuracy: 0.8710 - val_loss: 0.4719 - val_sparse_categorical_accuracy: 0.8493 - 722ms/epoch - 2ms/step

Epoch 52/100
469/469 - 1s - loss: 0.3642 - sparse_categorical_accuracy: 0.8709 - val_loss: 0.5226 - val_sparse_categorical_accuracy: 0.8308 - 761ms/epoch - 2ms/step

Epoch 53/100
469/469 - 1s - loss: 0.3601 - sparse_categorical_accuracy: 0.8731 - val_loss: 0.5038 - val_sparse_categorical_accuracy: 0.8375 - 733ms/epoch - 2ms/step

Epoch 54/100
469/469 - 1s - loss: 0.3578 - sparse_categorical_accuracy: 0.8733 - val_loss: 0.5036 - val_sparse_categorical_accuracy: 0.8395 - 732ms/epoch - 2ms/step

Epoch 55/100
469/469 - 1s - loss: 0.3587 - sparse_categorical_accuracy: 0.8724 - val_loss: 0.4898 - val_sparse_categorical_accuracy: 0.8467 - 788ms/epoch - 2ms/step

Epoch 56/100
469/469 - 1s - loss: 0.3560 - sparse_categorical_accuracy: 0.8737 - val_loss: 0.4699 - val_sparse_categorical_accuracy: 0.8474 - 756ms/epoch - 2ms/step

Epoch 57/100
469/469 - 1s - loss: 0.3545 - sparse_categorical_accuracy: 0.8738 - val_loss: 0.4842 - val_sparse_categorical_accuracy: 0.8407 - 717ms/epoch - 2ms/step

Epoch 58/100
469/469 - 1s - loss: 0.3566 - sparse_categorical_accuracy: 0.8736 - val_loss: 0.4989 - val_sparse_categorical_accuracy: 0.8407 - 717ms/epoch - 2ms/step

tegorical accuracy: 0.8452 - 807ms/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.3532 - sparse_categorical_accuracy: 0.8750 - val_loss: 0.4720 - val_sparse_ca
tegorical accuracy: 0.8455 - 808ms/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.3480 - sparse_categorical_accuracy: 0.8755 - val_loss: 0.4819 - val_sparse_ca
tegorical accuracy: 0.8453 - 806ms/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.3567 - sparse_categorical_accuracy: 0.8738 - val_loss: 0.4955 - val_sparse_ca
tegorical accuracy: 0.8408 - 693ms/epoch - 1ms/step
Epoch 62/100
469/469 - 1s - loss: 0.3534 - sparse_categorical_accuracy: 0.8747 - val_loss: 0.5261 - val_sparse_ca
tegorical accuracy: 0.8312 - 806ms/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 0.3524 - sparse_categorical_accuracy: 0.8758 - val_loss: 0.4638 - val_sparse_ca
tegorical accuracy: 0.8507 - 731ms/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.3488 - sparse_categorical_accuracy: 0.8747 - val_loss: 0.4932 - val_sparse_ca
tegorical accuracy: 0.8441 - 724ms/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.3477 - sparse_categorical_accuracy: 0.8766 - val_loss: 0.4903 - val_sparse_ca
tegorical accuracy: 0.8405 - 757ms/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.3522 - sparse_categorical_accuracy: 0.8752 - val_loss: 0.4797 - val_sparse_ca
tegorical accuracy: 0.8444 - 717ms/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.3456 - sparse_categorical_accuracy: 0.8760 - val_loss: 0.4921 - val_sparse_ca
tegorical accuracy: 0.8403 - 780ms/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 0.3457 - sparse_categorical_accuracy: 0.8770 - val_loss: 0.4966 - val_sparse_ca
tegorical accuracy: 0.8470 - 763ms/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 0.3454 - sparse_categorical_accuracy: 0.8771 - val_loss: 0.4803 - val_sparse_ca
tegorical accuracy: 0.8495 - 801ms/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.3479 - sparse_categorical_accuracy: 0.8763 - val_loss: 0.4792 - val_sparse_ca
tegorical accuracy: 0.8457 - 813ms/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.3448 - sparse_categorical_accuracy: 0.8764 - val_loss: 0.4963 - val_sparse_ca
tegorical accuracy: 0.8431 - 768ms/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.3441 - sparse_categorical_accuracy: 0.8773 - val_loss: 0.4930 - val_sparse_ca
tegorical accuracy: 0.8403 - 744ms/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.3394 - sparse_categorical_accuracy: 0.8777 - val_loss: 0.4775 - val_sparse_ca
tegorical accuracy: 0.8473 - 825ms/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.3390 - sparse_categorical_accuracy: 0.8804 - val_loss: 0.4920 - val_sparse_ca
tegorical accuracy: 0.8490 - 822ms/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.3433 - sparse_categorical_accuracy: 0.8787 - val_loss: 0.5049 - val_sparse_ca
tegorical accuracy: 0.8459 - 855ms/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.3428 - sparse_categorical_accuracy: 0.8770 - val_loss: 0.5065 - val_sparse_ca
tegorical accuracy: 0.8488 - 843ms/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.3386 - sparse_categorical_accuracy: 0.8790 - val_loss: 0.4846 - val_sparse_ca
tegorical accuracy: 0.8489 - 794ms/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.3441 - sparse_categorical_accuracy: 0.8783 - val_loss: 0.4929 - val_sparse_ca
tegorical accuracy: 0.8442 - 786ms/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.3418 - sparse_categorical_accuracy: 0.8780 - val_loss: 0.5024 - val_sparse_ca
tegorical accuracy: 0.8454 - 847ms/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.3380 - sparse_categorical_accuracy: 0.8796 - val_loss: 0.5419 - val_sparse_ca
tegorical accuracy: 0.8356 - 784ms/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.3393 - sparse_categorical_accuracy: 0.8788 - val_loss: 0.5143 - val_sparse_ca
tegorical accuracy: 0.8408 - 848ms/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.3382 - sparse_categorical_accuracy: 0.8790 - val_loss: 0.4885 - val_sparse_ca
tegorical accuracy: 0.8478 - 853ms/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.3363 - sparse_categorical_accuracy: 0.8802 - val_loss: 0.5088 - val_sparse_ca
tegorical accuracy: 0.8428 - 760ms/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.3402 - sparse_categorical_accuracy: 0.8784 - val_loss: 0.5034 - val_sparse_ca
tegorical accuracy: 0.8456 - 818ms/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.3359 - sparse_categorical_accuracy: 0.8794 - val_loss: 0.5041 - val_sparse_ca
tegorical accuracy: 0.8452 - 807ms/epoch - 2ms/step
Epoch 86/100

469/469 - 1s - loss: 0.3324 - sparse_categorical_accuracy: 0.8798 - val_loss: 0.5156 - val_sparse_categorical_accuracy: 0.8400 - 792ms/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.3446 - sparse_categorical_accuracy: 0.8775 - val_loss: 0.4815 - val_sparse_categorical_accuracy: 0.8492 - 737ms/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.3343 - sparse_categorical_accuracy: 0.8792 - val_loss: 0.5050 - val_sparse_categorical_accuracy: 0.8483 - 724ms/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.3318 - sparse_categorical_accuracy: 0.8811 - val_loss: 0.5097 - val_sparse_categorical_accuracy: 0.8409 - 746ms/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.3404 - sparse_categorical_accuracy: 0.8784 - val_loss: 0.5183 - val_sparse_categorical_accuracy: 0.8388 - 726ms/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.3370 - sparse_categorical_accuracy: 0.8802 - val_loss: 0.5189 - val_sparse_categorical_accuracy: 0.8476 - 731ms/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.3426 - sparse_categorical_accuracy: 0.8795 - val_loss: 0.5048 - val_sparse_categorical_accuracy: 0.8437 - 746ms/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.3325 - sparse_categorical_accuracy: 0.8817 - val_loss: 0.5129 - val_sparse_categorical_accuracy: 0.8462 - 786ms/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.3360 - sparse_categorical_accuracy: 0.8801 - val_loss: 0.5039 - val_sparse_categorical_accuracy: 0.8490 - 787ms/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.3324 - sparse_categorical_accuracy: 0.8809 - val_loss: 0.5192 - val_sparse_categorical_accuracy: 0.8419 - 772ms/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.3369 - sparse_categorical_accuracy: 0.8799 - val_loss: 0.5226 - val_sparse_categorical_accuracy: 0.8451 - 825ms/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.3339 - sparse_categorical_accuracy: 0.8790 - val_loss: 0.5168 - val_sparse_categorical_accuracy: 0.8479 - 735ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.3317 - sparse_categorical_accuracy: 0.8799 - val_loss: 0.5239 - val_sparse_categorical_accuracy: 0.8392 - 813ms/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.3356 - sparse_categorical_accuracy: 0.8803 - val_loss: 0.5009 - val_sparse_categorical_accuracy: 0.8443 - 748ms/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.3355 - sparse_categorical_accuracy: 0.8791 - val_loss: 0.5186 - val_sparse_categorical_accuracy: 0.8462 - 789ms/epoch - 2ms/step

In [10]:

```
model_kernel1.evaluate(x_test, y_test, verbose=2)
```

313/313 - 0s - loss: 0.5186 - sparse_categorical_accuracy: 0.8462 - 236ms/epoch - 755us/step

Out[10]:

```
[0.518607497215271, 0.8461999893188477]
```

Model with Ones kernel initializer

In [11]:

```
model_kernel2 = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(16, activation='relu', kernel_initializer='Ones'),
    tf.keras.layers.Dense(16, activation='relu', kernel_initializer='Ones'),
    tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='Ones')
])
```

In [12]:

```
model_kernel2.compile(optimizer='adam',
    loss='sparse_categorical_crossentropy',
    metrics=['sparse_categorical_accuracy'])
```

In [13]:

```
history_kernel2 = model_kernel2.fit(x_train, y_train,
    batch_size=128,
    epochs=100,
    validation_data=(x_test, y_test),
    verbose=2)
```

Epoch 1/100

469/469 - 1s - loss: 6881.9380 - sparse_categorical_accuracy: 0.1002 - val_loss: 2990.8706 - val_sparse_categorical_accuracy: 0.1000 - 1s/epoch - 3ms/step
Epoch 2/100
469/469 - 1s - loss: 2224.5042 - sparse_categorical_accuracy: 0.1009 - val_loss: 946.8256 - val_sparse_categorical_accuracy: 0.1000 - 805ms/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 767.5317 - sparse_categorical_accuracy: 0.0999 - val_loss: 422.9027 - val_sparse_categorical_accuracy: 0.1000 - 804ms/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 279.7441 - sparse_categorical_accuracy: 0.0996 - val_loss: 220.6814 - val_sparse_categorical_accuracy: 0.1000 - 738ms/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 107.6115 - sparse_categorical_accuracy: 0.0990 - val_loss: 79.8696 - val_sparse_categorical_accuracy: 0.1000 - 756ms/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 44.5764 - sparse_categorical_accuracy: 0.0997 - val_loss: 13.1308 - val_sparse_categorical_accuracy: 0.1115 - 702ms/epoch - 1ms/step
Epoch 7/100
469/469 - 1s - loss: 19.4519 - sparse_categorical_accuracy: 0.1028 - val_loss: 13.1924 - val_sparse_categorical_accuracy: 0.1000 - 762ms/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 6.9886 - sparse_categorical_accuracy: 0.1190 - val_loss: 2.6603 - val_sparse_categorical_accuracy: 0.1714 - 711ms/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 3.2249 - sparse_categorical_accuracy: 0.1546 - val_loss: 4.5615 - val_sparse_categorical_accuracy: 0.1259 - 703ms/epoch - 1ms/step
Epoch 10/100
469/469 - 1s - loss: 3.1173 - sparse_categorical_accuracy: 0.1592 - val_loss: 2.3968 - val_sparse_categorical_accuracy: 0.1661 - 794ms/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 2.7836 - sparse_categorical_accuracy: 0.1721 - val_loss: 2.3883 - val_sparse_categorical_accuracy: 0.2112 - 810ms/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 2.4630 - sparse_categorical_accuracy: 0.1869 - val_loss: 2.2435 - val_sparse_categorical_accuracy: 0.2082 - 740ms/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 2.3399 - sparse_categorical_accuracy: 0.1942 - val_loss: 2.1812 - val_sparse_categorical_accuracy: 0.2213 - 726ms/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 2.1781 - sparse_categorical_accuracy: 0.2027 - val_loss: 2.2672 - val_sparse_categorical_accuracy: 0.1978 - 825ms/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 2.1228 - sparse_categorical_accuracy: 0.2035 - val_loss: 2.1699 - val_sparse_categorical_accuracy: 0.1839 - 718ms/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 2.0298 - sparse_categorical_accuracy: 0.2140 - val_loss: 2.0174 - val_sparse_categorical_accuracy: 0.2021 - 786ms/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 1.9634 - sparse_categorical_accuracy: 0.2188 - val_loss: 1.9316 - val_sparse_categorical_accuracy: 0.2066 - 738ms/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 1.9076 - sparse_categorical_accuracy: 0.2262 - val_loss: 1.8743 - val_sparse_categorical_accuracy: 0.2419 - 839ms/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 1.8551 - sparse_categorical_accuracy: 0.2419 - val_loss: 1.8210 - val_sparse_categorical_accuracy: 0.2619 - 819ms/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 1.7980 - sparse_categorical_accuracy: 0.2718 - val_loss: 1.7748 - val_sparse_categorical_accuracy: 0.2788 - 795ms/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 1.7493 - sparse_categorical_accuracy: 0.2913 - val_loss: 1.7168 - val_sparse_categorical_accuracy: 0.2822 - 743ms/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 1.7072 - sparse_categorical_accuracy: 0.3017 - val_loss: 1.6729 - val_sparse_categorical_accuracy: 0.3288 - 724ms/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 1.6664 - sparse_categorical_accuracy: 0.3130 - val_loss: 1.6381 - val_sparse_categorical_accuracy: 0.3080 - 773ms/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 1.6296 - sparse_categorical_accuracy: 0.3284 - val_loss: 1.6062 - val_sparse_categorical_accuracy: 0.3273 - 814ms/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 1.5981 - sparse_categorical_accuracy: 0.3424 - val_loss: 1.6072 - val_sparse_categorical_accuracy: 0.3079 - 1s/epoch - 3ms/step
Epoch 26/100
469/469 - 1s - loss: 1.5683 - sparse_categorical_accuracy: 0.3489 - val_loss: 1.5438 - val_sparse_categorical_accuracy: 0.3551 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 1.5408 - sparse_categorical_accuracy: 0.3604 - val_loss: 1.5242 - val_sparse_categorical_accuracy: 0.3623 - 840ms/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 1.5169 - sparse_categorical_accuracy: 0.3672 - val_loss: 1.5300 - val_sparse_categorical_accuracy: 0.3421 - 759ms/epoch - 2ms/step

Epoch 29/100
469/469 - 1s - loss: 1.4998 - sparse_categorical_accuracy: 0.3693 - val_loss: 1.4837 - val_sparse_categorical_accuracy: 0.3652 - 808ms/epoch - 2ms/step

Epoch 30/100
469/469 - 1s - loss: 1.4851 - sparse_categorical_accuracy: 0.3733 - val_loss: 1.5108 - val_sparse_categorical_accuracy: 0.3654 - 803ms/epoch - 2ms/step

Epoch 31/100
469/469 - 1s - loss: 1.4819 - sparse_categorical_accuracy: 0.3659 - val_loss: 1.4769 - val_sparse_categorical_accuracy: 0.3554 - 783ms/epoch - 2ms/step

Epoch 32/100
469/469 - 1s - loss: 1.4612 - sparse_categorical_accuracy: 0.3553 - val_loss: 1.4408 - val_sparse_categorical_accuracy: 0.3676 - 742ms/epoch - 2ms/step

Epoch 33/100
469/469 - 1s - loss: 1.4503 - sparse_categorical_accuracy: 0.3565 - val_loss: 1.4471 - val_sparse_categorical_accuracy: 0.3563 - 755ms/epoch - 2ms/step

Epoch 34/100
469/469 - 1s - loss: 1.4413 - sparse_categorical_accuracy: 0.3585 - val_loss: 1.4258 - val_sparse_categorical_accuracy: 0.3665 - 802ms/epoch - 2ms/step

Epoch 35/100
469/469 - 1s - loss: 1.4423 - sparse_categorical_accuracy: 0.3608 - val_loss: 1.4517 - val_sparse_categorical_accuracy: 0.3413 - 728ms/epoch - 2ms/step

Epoch 36/100
469/469 - 1s - loss: 1.4352 - sparse_categorical_accuracy: 0.3597 - val_loss: 1.4664 - val_sparse_categorical_accuracy: 0.3299 - 781ms/epoch - 2ms/step

Epoch 37/100
469/469 - 1s - loss: 1.4351 - sparse_categorical_accuracy: 0.3594 - val_loss: 1.4500 - val_sparse_categorical_accuracy: 0.3491 - 733ms/epoch - 2ms/step

Epoch 38/100
469/469 - 1s - loss: 1.4253 - sparse_categorical_accuracy: 0.3637 - val_loss: 1.4202 - val_sparse_categorical_accuracy: 0.3647 - 698ms/epoch - 1ms/step

Epoch 39/100
469/469 - 1s - loss: 1.4300 - sparse_categorical_accuracy: 0.3609 - val_loss: 1.4315 - val_sparse_categorical_accuracy: 0.3753 - 794ms/epoch - 2ms/step

Epoch 40/100
469/469 - 1s - loss: 1.4212 - sparse_categorical_accuracy: 0.3625 - val_loss: 1.4545 - val_sparse_categorical_accuracy: 0.3594 - 830ms/epoch - 2ms/step

Epoch 41/100
469/469 - 1s - loss: 1.4229 - sparse_categorical_accuracy: 0.3626 - val_loss: 1.4247 - val_sparse_categorical_accuracy: 0.3695 - 749ms/epoch - 2ms/step

Epoch 42/100
469/469 - 1s - loss: 1.4179 - sparse_categorical_accuracy: 0.3657 - val_loss: 1.4496 - val_sparse_categorical_accuracy: 0.3513 - 813ms/epoch - 2ms/step

Epoch 43/100
469/469 - 1s - loss: 1.4140 - sparse_categorical_accuracy: 0.3659 - val_loss: 1.4889 - val_sparse_categorical_accuracy: 0.3288 - 813ms/epoch - 2ms/step

Epoch 44/100
469/469 - 1s - loss: 1.4140 - sparse_categorical_accuracy: 0.3614 - val_loss: 1.4149 - val_sparse_categorical_accuracy: 0.3656 - 797ms/epoch - 2ms/step

Epoch 45/100
469/469 - 1s - loss: 1.4113 - sparse_categorical_accuracy: 0.3660 - val_loss: 1.4365 - val_sparse_categorical_accuracy: 0.3565 - 794ms/epoch - 2ms/step

Epoch 46/100
469/469 - 1s - loss: 1.4097 - sparse_categorical_accuracy: 0.3650 - val_loss: 1.4222 - val_sparse_categorical_accuracy: 0.3598 - 726ms/epoch - 2ms/step

Epoch 47/100
469/469 - 1s - loss: 1.4099 - sparse_categorical_accuracy: 0.3652 - val_loss: 1.4613 - val_sparse_categorical_accuracy: 0.3580 - 771ms/epoch - 2ms/step

Epoch 48/100
469/469 - 1s - loss: 1.4124 - sparse_categorical_accuracy: 0.3670 - val_loss: 1.4249 - val_sparse_categorical_accuracy: 0.3561 - 738ms/epoch - 2ms/step

Epoch 49/100
469/469 - 1s - loss: 1.4073 - sparse_categorical_accuracy: 0.3640 - val_loss: 1.4019 - val_sparse_categorical_accuracy: 0.3718 - 721ms/epoch - 2ms/step

Epoch 50/100
469/469 - 1s - loss: 1.4063 - sparse_categorical_accuracy: 0.3646 - val_loss: 1.4449 - val_sparse_categorical_accuracy: 0.3696 - 775ms/epoch - 2ms/step

Epoch 51/100
469/469 - 1s - loss: 1.3991 - sparse_categorical_accuracy: 0.3690 - val_loss: 1.4068 - val_sparse_categorical_accuracy: 0.3721 - 739ms/epoch - 2ms/step

Epoch 52/100
469/469 - 1s - loss: 1.4015 - sparse_categorical_accuracy: 0.3656 - val_loss: 1.4092 - val_sparse_categorical_accuracy: 0.3485 - 803ms/epoch - 2ms/step

Epoch 53/100
469/469 - 1s - loss: 1.3981 - sparse_categorical_accuracy: 0.3682 - val_loss: 1.4336 - val_sparse_categorical_accuracy: 0.3826 - 791ms/epoch - 2ms/step

Epoch 54/100
469/469 - 1s - loss: 1.4008 - sparse_categorical_accuracy: 0.3672 - val_loss: 1.3954 - val_sparse_categorical_accuracy: 0.3680 - 727ms/epoch - 2ms/step

Epoch 55/100
469/469 - 1s - loss: 1.3970 - sparse_categorical_accuracy: 0.3673 - val_loss: 1.4134 - val_sparse_categorical_accuracy: 0.3778 - 834ms/epoch - 2ms/step

Epoch 56/100
469/469 - 1s - loss: 1.3918 - sparse_categorical_accuracy: 0.3712 - val_loss: 1.4030 - val_sparse_categorical_accuracy: 0.3712 - 791ms/epoch - 2ms/step

tegorical_accuracy: 0.3669 - 765ms/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 1.3933 - sparse_categorical_accuracy: 0.3658 - val_loss: 1.4188 - val_sparse_ca
tegorical_accuracy: 0.3729 - 865ms/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 1.3855 - sparse_categorical_accuracy: 0.3695 - val_loss: 1.4285 - val_sparse_ca
tegorical_accuracy: 0.3592 - 755ms/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 1.3950 - sparse_categorical_accuracy: 0.3651 - val_loss: 1.4083 - val_sparse_ca
tegorical_accuracy: 0.3593 - 739ms/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 1.3908 - sparse_categorical_accuracy: 0.3659 - val_loss: 1.4052 - val_sparse_ca
tegorical_accuracy: 0.3647 - 817ms/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 1.3861 - sparse_categorical_accuracy: 0.3681 - val_loss: 1.4046 - val_sparse_ca
tegorical_accuracy: 0.3920 - 695ms/epoch - 1ms/step
Epoch 62/100
469/469 - 1s - loss: 1.3875 - sparse_categorical_accuracy: 0.3689 - val_loss: 1.4020 - val_sparse_ca
tegorical_accuracy: 0.3668 - 747ms/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 1.3850 - sparse_categorical_accuracy: 0.3687 - val_loss: 1.4030 - val_sparse_ca
tegorical_accuracy: 0.3699 - 789ms/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 1.3850 - sparse_categorical_accuracy: 0.3696 - val_loss: 1.4743 - val_sparse_ca
tegorical_accuracy: 0.3223 - 792ms/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 1.3850 - sparse_categorical_accuracy: 0.3677 - val_loss: 1.4417 - val_sparse_ca
tegorical_accuracy: 0.3512 - 742ms/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 1.3873 - sparse_categorical_accuracy: 0.3660 - val_loss: 1.4008 - val_sparse_ca
tegorical_accuracy: 0.3731 - 796ms/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 1.3829 - sparse_categorical_accuracy: 0.3669 - val_loss: 1.4142 - val_sparse_ca
tegorical_accuracy: 0.3562 - 738ms/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 1.3824 - sparse_categorical_accuracy: 0.3686 - val_loss: 1.3990 - val_sparse_ca
tegorical_accuracy: 0.3491 - 779ms/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 1.3838 - sparse_categorical_accuracy: 0.3706 - val_loss: 1.4435 - val_sparse_ca
tegorical_accuracy: 0.3672 - 740ms/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 1.3824 - sparse_categorical_accuracy: 0.3688 - val_loss: 1.3961 - val_sparse_ca
tegorical_accuracy: 0.3848 - 747ms/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 1.3824 - sparse_categorical_accuracy: 0.3682 - val_loss: 1.4107 - val_sparse_ca
tegorical_accuracy: 0.3606 - 713ms/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 1.3791 - sparse_categorical_accuracy: 0.3670 - val_loss: 1.4289 - val_sparse_ca
tegorical_accuracy: 0.3574 - 727ms/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 1.3777 - sparse_categorical_accuracy: 0.3719 - val_loss: 1.3871 - val_sparse_ca
tegorical_accuracy: 0.3725 - 802ms/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 1.3740 - sparse_categorical_accuracy: 0.3716 - val_loss: 1.4123 - val_sparse_ca
tegorical_accuracy: 0.3695 - 744ms/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 1.3758 - sparse_categorical_accuracy: 0.3705 - val_loss: 1.3942 - val_sparse_ca
tegorical_accuracy: 0.3757 - 825ms/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 1.3750 - sparse_categorical_accuracy: 0.3722 - val_loss: 1.5174 - val_sparse_ca
tegorical_accuracy: 0.3208 - 788ms/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 1.3800 - sparse_categorical_accuracy: 0.3675 - val_loss: 1.4087 - val_sparse_ca
tegorical_accuracy: 0.3766 - 782ms/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 1.3788 - sparse_categorical_accuracy: 0.3696 - val_loss: 1.4082 - val_sparse_ca
tegorical_accuracy: 0.3757 - 799ms/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 1.3759 - sparse_categorical_accuracy: 0.3701 - val_loss: 1.4018 - val_sparse_ca
tegorical_accuracy: 0.3864 - 805ms/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 1.3807 - sparse_categorical_accuracy: 0.3695 - val_loss: 1.4035 - val_sparse_ca
tegorical_accuracy: 0.3668 - 731ms/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 1.3805 - sparse_categorical_accuracy: 0.3701 - val_loss: 1.4380 - val_sparse_ca
tegorical_accuracy: 0.3591 - 720ms/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 1.3757 - sparse_categorical_accuracy: 0.3753 - val_loss: 1.4258 - val_sparse_ca
tegorical_accuracy: 0.3624 - 811ms/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 1.3755 - sparse_categorical_accuracy: 0.3711 - val_loss: 1.4033 - val_sparse_ca
tegorical_accuracy: 0.3669 - 777ms/epoch - 2ms/step
Epoch 84/100

```

469/469 - 1s - loss: 1.3787 - sparse_categorical_accuracy: 0.3693 - val_loss: 1.4057 - val_sparse_categorical_accuracy: 0.3582 - 712ms/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 1.3743 - sparse_categorical_accuracy: 0.3744 - val_loss: 1.4219 - val_sparse_categorical_accuracy: 0.3530 - 718ms/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 1.3729 - sparse_categorical_accuracy: 0.3761 - val_loss: 1.4394 - val_sparse_categorical_accuracy: 0.3582 - 771ms/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 1.3741 - sparse_categorical_accuracy: 0.3729 - val_loss: 1.4225 - val_sparse_categorical_accuracy: 0.3508 - 738ms/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 1.3777 - sparse_categorical_accuracy: 0.3716 - val_loss: 1.4336 - val_sparse_categorical_accuracy: 0.3625 - 837ms/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 1.3821 - sparse_categorical_accuracy: 0.3690 - val_loss: 1.4470 - val_sparse_categorical_accuracy: 0.3698 - 796ms/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 1.3758 - sparse_categorical_accuracy: 0.3701 - val_loss: 1.3985 - val_sparse_categorical_accuracy: 0.3844 - 826ms/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 1.3708 - sparse_categorical_accuracy: 0.3734 - val_loss: 1.4024 - val_sparse_categorical_accuracy: 0.3713 - 830ms/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 1.3800 - sparse_categorical_accuracy: 0.3695 - val_loss: 1.4444 - val_sparse_categorical_accuracy: 0.3624 - 818ms/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 1.3710 - sparse_categorical_accuracy: 0.3747 - val_loss: 1.4035 - val_sparse_categorical_accuracy: 0.3743 - 713ms/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 1.3775 - sparse_categorical_accuracy: 0.3699 - val_loss: 1.4497 - val_sparse_categorical_accuracy: 0.3680 - 800ms/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 1.3715 - sparse_categorical_accuracy: 0.3756 - val_loss: 1.4144 - val_sparse_categorical_accuracy: 0.3691 - 824ms/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 1.3761 - sparse_categorical_accuracy: 0.3738 - val_loss: 1.4591 - val_sparse_categorical_accuracy: 0.3552 - 814ms/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 1.3715 - sparse_categorical_accuracy: 0.3760 - val_loss: 1.4292 - val_sparse_categorical_accuracy: 0.3640 - 728ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 1.3732 - sparse_categorical_accuracy: 0.3724 - val_loss: 1.4093 - val_sparse_categorical_accuracy: 0.3683 - 775ms/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 1.3661 - sparse_categorical_accuracy: 0.3783 - val_loss: 1.4363 - val_sparse_categorical_accuracy: 0.3576 - 707ms/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 1.3787 - sparse_categorical_accuracy: 0.3703 - val_loss: 1.4232 - val_sparse_categorical_accuracy: 0.3733 - 777ms/epoch - 2ms/step

```

In [14]:

```
model_kernel2.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 1.4232 - sparse_categorical_accuracy: 0.3733 - 235ms/epoch - 751us/step
```

Out[14]:

```
[1.423214316368103, 0.3732999861240387]
```

Model with random_uniform initializer

In [15]:

```

model_kernel3 = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
    tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
    tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])

```

In [16]:

```

model_kernel3.compile(optimizer='adam',
    loss='sparse_categorical_crossentropy',
    metrics=['sparse_categorical_accuracy']
)

```

In [17]:

```
history_kernel3 = model_kernel3.fit(x_train, y_train,  
                                     batch_size=128,  
                                     epochs=100,  
                                     validation_data=(x_test, y_test),  
                                     verbose=2  
                                     )
```

Epoch 1/100

469/469 - 1s - loss: 0.6256 - sparse_categorical_accuracy: 0.7747 - val_loss: 0.4890 - val_sparse_categorical_accuracy: 0.8264 - 1s/epoch - 3ms/step

Epoch 2/100

469/469 - 1s - loss: 0.4379 - sparse_categorical_accuracy: 0.8411 - val_loss: 0.4503 - val_sparse_categorical_accuracy: 0.8392 - 734ms/epoch - 2ms/step

Epoch 3/100

469/469 - 1s - loss: 0.4056 - sparse_categorical_accuracy: 0.8547 - val_loss: 0.4454 - val_sparse_categorical_accuracy: 0.8427 - 827ms/epoch - 2ms/step

Epoch 4/100

469/469 - 1s - loss: 0.3872 - sparse_categorical_accuracy: 0.8615 - val_loss: 0.4380 - val_sparse_categorical_accuracy: 0.8460 - 748ms/epoch - 2ms/step

Epoch 5/100

469/469 - 1s - loss: 0.3771 - sparse_categorical_accuracy: 0.8625 - val_loss: 0.4272 - val_sparse_categorical_accuracy: 0.8513 - 808ms/epoch - 2ms/step

Epoch 6/100

469/469 - 1s - loss: 0.3647 - sparse_categorical_accuracy: 0.8677 - val_loss: 0.4168 - val_sparse_categorical_accuracy: 0.8520 - 827ms/epoch - 2ms/step

Epoch 7/100

469/469 - 1s - loss: 0.3537 - sparse_categorical_accuracy: 0.8718 - val_loss: 0.4171 - val_sparse_categorical_accuracy: 0.8537 - 741ms/epoch - 2ms/step

Epoch 8/100

469/469 - 1s - loss: 0.3530 - sparse_categorical_accuracy: 0.8719 - val_loss: 0.4360 - val_sparse_categorical_accuracy: 0.8430 - 795ms/epoch - 2ms/step

Epoch 9/100

469/469 - 1s - loss: 0.3427 - sparse_categorical_accuracy: 0.8753 - val_loss: 0.4447 - val_sparse_categorical_accuracy: 0.8466 - 764ms/epoch - 2ms/step

Epoch 10/100

469/469 - 1s - loss: 0.3388 - sparse_categorical_accuracy: 0.8763 - val_loss: 0.4157 - val_sparse_categorical_accuracy: 0.8600 - 782ms/epoch - 2ms/step

Epoch 11/100

469/469 - 1s - loss: 0.3337 - sparse_categorical_accuracy: 0.8789 - val_loss: 0.4066 - val_sparse_categorical_accuracy: 0.8613 - 719ms/epoch - 2ms/step

Epoch 12/100

469/469 - 1s - loss: 0.3255 - sparse_categorical_accuracy: 0.8807 - val_loss: 0.4065 - val_sparse_categorical_accuracy: 0.8623 - 747ms/epoch - 2ms/step

Epoch 13/100

469/469 - 1s - loss: 0.3244 - sparse_categorical_accuracy: 0.8823 - val_loss: 0.4161 - val_sparse_categorical_accuracy: 0.8557 - 721ms/epoch - 2ms/step

Epoch 14/100

469/469 - 1s - loss: 0.3260 - sparse_categorical_accuracy: 0.8814 - val_loss: 0.4164 - val_sparse_categorical_accuracy: 0.8562 - 808ms/epoch - 2ms/step

Epoch 15/100

469/469 - 1s - loss: 0.3209 - sparse_categorical_accuracy: 0.8832 - val_loss: 0.4165 - val_sparse_categorical_accuracy: 0.8529 - 724ms/epoch - 2ms/step

Epoch 16/100

469/469 - 1s - loss: 0.3194 - sparse_categorical_accuracy: 0.8826 - val_loss: 0.4412 - val_sparse_categorical_accuracy: 0.8514 - 758ms/epoch - 2ms/step

Epoch 17/100

469/469 - 1s - loss: 0.3125 - sparse_categorical_accuracy: 0.8863 - val_loss: 0.4106 - val_sparse_categorical_accuracy: 0.8609 - 792ms/epoch - 2ms/step

Epoch 18/100

469/469 - 1s - loss: 0.3130 - sparse_categorical_accuracy: 0.8842 - val_loss: 0.4144 - val_sparse_categorical_accuracy: 0.8586 - 813ms/epoch - 2ms/step

Epoch 19/100

469/469 - 1s - loss: 0.3147 - sparse_categorical_accuracy: 0.8838 - val_loss: 0.4173 - val_sparse_categorical_accuracy: 0.8600 - 788ms/epoch - 2ms/step

Epoch 20/100

469/469 - 1s - loss: 0.3052 - sparse_categorical_accuracy: 0.8873 - val_loss: 0.4448 - val_sparse_categorical_accuracy: 0.8533 - 730ms/epoch - 2ms/step

Epoch 21/100

469/469 - 1s - loss: 0.3072 - sparse_categorical_accuracy: 0.8862 - val_loss: 0.4201 - val_sparse_categorical_accuracy: 0.8626 - 687ms/epoch - 1ms/step

Epoch 22/100

469/469 - 1s - loss: 0.3060 - sparse_categorical_accuracy: 0.8867 - val_loss: 0.4314 - val_sparse_categorical_accuracy: 0.8640 - 789ms/epoch - 2ms/step

Epoch 23/100

469/469 - 1s - loss: 0.3049 - sparse_categorical_accuracy: 0.8873 - val_loss: 0.4295 - val_sparse_categorical_accuracy: 0.8568 - 803ms/epoch - 2ms/step

Epoch 24/100

469/469 - 1s - loss: 0.3028 - sparse_categorical_accuracy: 0.8894 - val_loss: 0.4395 - val_sparse_categorical_accuracy: 0.8550 - 786ms/epoch - 2ms/step

Epoch 25/100

469/469 - 1s - loss: 0.2973 - sparse_categorical_accuracy: 0.8911 - val_loss: 0.4091 - val_sparse_categorical_accuracy: 0.8550 - 786ms/epoch - 2ms/step

tegorical accuracy: 0.8622 - 725ms/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.2993 - sparse_categorical_accuracy: 0.8893 - val_loss: 0.4190 - val_sparse_categorical_accuracy: 0.8586 - 724ms/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.2925 - sparse_categorical_accuracy: 0.8924 - val_loss: 0.4253 - val_sparse_categorical_accuracy: 0.8611 - 806ms/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.2968 - sparse_categorical_accuracy: 0.8905 - val_loss: 0.4438 - val_sparse_categorical_accuracy: 0.8558 - 741ms/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.2952 - sparse_categorical_accuracy: 0.8917 - val_loss: 0.4138 - val_sparse_categorical_accuracy: 0.8618 - 780ms/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.2935 - sparse_categorical_accuracy: 0.8928 - val_loss: 0.4106 - val_sparse_categorical_accuracy: 0.8622 - 743ms/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.2888 - sparse_categorical_accuracy: 0.8931 - val_loss: 0.4111 - val_sparse_categorical_accuracy: 0.8636 - 808ms/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.2884 - sparse_categorical_accuracy: 0.8939 - val_loss: 0.4319 - val_sparse_categorical_accuracy: 0.8608 - 832ms/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.2894 - sparse_categorical_accuracy: 0.8929 - val_loss: 0.4470 - val_sparse_categorical_accuracy: 0.8542 - 789ms/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.2909 - sparse_categorical_accuracy: 0.8921 - val_loss: 0.4393 - val_sparse_categorical_accuracy: 0.8596 - 786ms/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.2865 - sparse_categorical_accuracy: 0.8935 - val_loss: 0.4485 - val_sparse_categorical_accuracy: 0.8539 - 711ms/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.2855 - sparse_categorical_accuracy: 0.8952 - val_loss: 0.4786 - val_sparse_categorical_accuracy: 0.8491 - 731ms/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.2843 - sparse_categorical_accuracy: 0.8952 - val_loss: 0.4487 - val_sparse_categorical_accuracy: 0.8591 - 777ms/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.2794 - sparse_categorical_accuracy: 0.8960 - val_loss: 0.4357 - val_sparse_categorical_accuracy: 0.8657 - 733ms/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.2815 - sparse_categorical_accuracy: 0.8953 - val_loss: 0.4563 - val_sparse_categorical_accuracy: 0.8586 - 795ms/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 0.2829 - sparse_categorical_accuracy: 0.8958 - val_loss: 0.4302 - val_sparse_categorical_accuracy: 0.8644 - 795ms/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.2789 - sparse_categorical_accuracy: 0.8977 - val_loss: 0.4610 - val_sparse_categorical_accuracy: 0.8559 - 803ms/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.2798 - sparse_categorical_accuracy: 0.8970 - val_loss: 0.4333 - val_sparse_categorical_accuracy: 0.8606 - 801ms/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.2814 - sparse_categorical_accuracy: 0.8946 - val_loss: 0.4441 - val_sparse_categorical_accuracy: 0.8626 - 800ms/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.2772 - sparse_categorical_accuracy: 0.8971 - val_loss: 0.4399 - val_sparse_categorical_accuracy: 0.8606 - 719ms/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.2751 - sparse_categorical_accuracy: 0.8986 - val_loss: 0.4446 - val_sparse_categorical_accuracy: 0.8652 - 855ms/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.2779 - sparse_categorical_accuracy: 0.8973 - val_loss: 0.4500 - val_sparse_categorical_accuracy: 0.8618 - 760ms/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.2802 - sparse_categorical_accuracy: 0.8961 - val_loss: 0.4609 - val_sparse_categorical_accuracy: 0.8554 - 742ms/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.2711 - sparse_categorical_accuracy: 0.8986 - val_loss: 0.4637 - val_sparse_categorical_accuracy: 0.8630 - 718ms/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.2733 - sparse_categorical_accuracy: 0.8978 - val_loss: 0.4409 - val_sparse_categorical_accuracy: 0.8598 - 789ms/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.2711 - sparse_categorical_accuracy: 0.8985 - val_loss: 0.4512 - val_sparse_categorical_accuracy: 0.8613 - 728ms/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.2681 - sparse_categorical_accuracy: 0.9007 - val_loss: 0.4572 - val_sparse_categorical_accuracy: 0.8663 - 721ms/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.2715 - sparse_categorical_accuracy: 0.8979 - val_loss: 0.4527 - val_sparse_categorical_accuracy: 0.8588 - 823ms/epoch - 2ms/step
Epoch 53/100

469/469 - 1s - loss: 0.2769 - sparse_categorical_accuracy: 0.8965 - val_loss: 0.4715 - val_sparse_categorical_accuracy: 0.8534 - 782ms/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.2734 - sparse_categorical_accuracy: 0.8978 - val_loss: 0.4691 - val_sparse_categorical_accuracy: 0.8638 - 785ms/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.2708 - sparse_categorical_accuracy: 0.8988 - val_loss: 0.5165 - val_sparse_categorical_accuracy: 0.8534 - 695ms/epoch - 1ms/step
Epoch 56/100
469/469 - 1s - loss: 0.2698 - sparse_categorical_accuracy: 0.8998 - val_loss: 0.4792 - val_sparse_categorical_accuracy: 0.8578 - 795ms/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 0.2704 - sparse_categorical_accuracy: 0.8989 - val_loss: 0.4577 - val_sparse_categorical_accuracy: 0.8581 - 824ms/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.2680 - sparse_categorical_accuracy: 0.8998 - val_loss: 0.4691 - val_sparse_categorical_accuracy: 0.8548 - 796ms/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.2674 - sparse_categorical_accuracy: 0.9000 - val_loss: 0.4716 - val_sparse_categorical_accuracy: 0.8636 - 727ms/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.2659 - sparse_categorical_accuracy: 0.9009 - val_loss: 0.4594 - val_sparse_categorical_accuracy: 0.8668 - 788ms/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.2703 - sparse_categorical_accuracy: 0.8991 - val_loss: 0.4684 - val_sparse_categorical_accuracy: 0.8636 - 812ms/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.2655 - sparse_categorical_accuracy: 0.9005 - val_loss: 0.4797 - val_sparse_categorical_accuracy: 0.8584 - 837ms/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 0.2639 - sparse_categorical_accuracy: 0.9007 - val_loss: 0.4757 - val_sparse_categorical_accuracy: 0.8632 - 773ms/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.2649 - sparse_categorical_accuracy: 0.9007 - val_loss: 0.4854 - val_sparse_categorical_accuracy: 0.8659 - 792ms/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.2670 - sparse_categorical_accuracy: 0.9000 - val_loss: 0.4800 - val_sparse_categorical_accuracy: 0.8607 - 807ms/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.2671 - sparse_categorical_accuracy: 0.8999 - val_loss: 0.4630 - val_sparse_categorical_accuracy: 0.8649 - 732ms/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.2663 - sparse_categorical_accuracy: 0.9002 - val_loss: 0.4912 - val_sparse_categorical_accuracy: 0.8549 - 719ms/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 0.2628 - sparse_categorical_accuracy: 0.9015 - val_loss: 0.4831 - val_sparse_categorical_accuracy: 0.8593 - 810ms/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 0.2625 - sparse_categorical_accuracy: 0.9017 - val_loss: 0.4869 - val_sparse_categorical_accuracy: 0.8632 - 728ms/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.2683 - sparse_categorical_accuracy: 0.8995 - val_loss: 0.5016 - val_sparse_categorical_accuracy: 0.8542 - 728ms/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.2666 - sparse_categorical_accuracy: 0.9007 - val_loss: 0.5048 - val_sparse_categorical_accuracy: 0.8622 - 726ms/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.2585 - sparse_categorical_accuracy: 0.9030 - val_loss: 0.4846 - val_sparse_categorical_accuracy: 0.8624 - 786ms/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.2637 - sparse_categorical_accuracy: 0.9014 - val_loss: 0.5099 - val_sparse_categorical_accuracy: 0.8607 - 827ms/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.2596 - sparse_categorical_accuracy: 0.9020 - val_loss: 0.4932 - val_sparse_categorical_accuracy: 0.8520 - 730ms/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.2607 - sparse_categorical_accuracy: 0.9019 - val_loss: 0.4968 - val_sparse_categorical_accuracy: 0.8610 - 784ms/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.2591 - sparse_categorical_accuracy: 0.9033 - val_loss: 0.5033 - val_sparse_categorical_accuracy: 0.8567 - 733ms/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.2630 - sparse_categorical_accuracy: 0.9011 - val_loss: 0.5240 - val_sparse_categorical_accuracy: 0.8604 - 740ms/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.2650 - sparse_categorical_accuracy: 0.9004 - val_loss: 0.4869 - val_sparse_categorical_accuracy: 0.8617 - 760ms/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.2601 - sparse_categorical_accuracy: 0.9027 - val_loss: 0.4988 - val_sparse_categorical_accuracy: 0.8615 - 709ms/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.2607 - sparse_categorical_accuracy: 0.9017 - val_loss: 0.5060 - val_sparse_categorical_accuracy: 0.8588 - 716ms/epoch - 2ms/step


```

Epoch 81/100
469/469 - 1s - loss: 0.2595 - sparse_categorical_accuracy: 0.9036 - val_loss: 0.5207 - val_sparse_categorical_accuracy: 0.8554 - 793ms/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.2607 - sparse_categorical_accuracy: 0.9025 - val_loss: 0.4945 - val_sparse_categorical_accuracy: 0.8571 - 833ms/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.2640 - sparse_categorical_accuracy: 0.9012 - val_loss: 0.5128 - val_sparse_categorical_accuracy: 0.8624 - 745ms/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.2582 - sparse_categorical_accuracy: 0.9038 - val_loss: 0.5228 - val_sparse_categorical_accuracy: 0.8577 - 801ms/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.2566 - sparse_categorical_accuracy: 0.9041 - val_loss: 0.4950 - val_sparse_categorical_accuracy: 0.8587 - 724ms/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.2557 - sparse_categorical_accuracy: 0.9048 - val_loss: 0.5050 - val_sparse_categorical_accuracy: 0.8609 - 837ms/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.2576 - sparse_categorical_accuracy: 0.9045 - val_loss: 0.5298 - val_sparse_categorical_accuracy: 0.8524 - 805ms/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.2529 - sparse_categorical_accuracy: 0.9053 - val_loss: 0.5161 - val_sparse_categorical_accuracy: 0.8610 - 812ms/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.2583 - sparse_categorical_accuracy: 0.9030 - val_loss: 0.5033 - val_sparse_categorical_accuracy: 0.8571 - 798ms/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.2565 - sparse_categorical_accuracy: 0.9042 - val_loss: 0.5232 - val_sparse_categorical_accuracy: 0.8559 - 750ms/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.2548 - sparse_categorical_accuracy: 0.9038 - val_loss: 0.5122 - val_sparse_categorical_accuracy: 0.8625 - 811ms/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.2589 - sparse_categorical_accuracy: 0.9033 - val_loss: 0.5192 - val_sparse_categorical_accuracy: 0.8513 - 757ms/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.2529 - sparse_categorical_accuracy: 0.9054 - val_loss: 0.5270 - val_sparse_categorical_accuracy: 0.8600 - 744ms/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.2560 - sparse_categorical_accuracy: 0.9043 - val_loss: 0.5262 - val_sparse_categorical_accuracy: 0.8558 - 751ms/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.2553 - sparse_categorical_accuracy: 0.9033 - val_loss: 0.5177 - val_sparse_categorical_accuracy: 0.8671 - 763ms/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.2512 - sparse_categorical_accuracy: 0.9064 - val_loss: 0.5198 - val_sparse_categorical_accuracy: 0.8624 - 797ms/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.2547 - sparse_categorical_accuracy: 0.9050 - val_loss: 0.5082 - val_sparse_categorical_accuracy: 0.8620 - 731ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.2565 - sparse_categorical_accuracy: 0.9030 - val_loss: 0.5348 - val_sparse_categorical_accuracy: 0.8618 - 777ms/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.2497 - sparse_categorical_accuracy: 0.9059 - val_loss: 0.5507 - val_sparse_categorical_accuracy: 0.8565 - 749ms/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.2541 - sparse_categorical_accuracy: 0.9050 - val_loss: 0.5425 - val_sparse_categorical_accuracy: 0.8538 - 738ms/epoch - 2ms/step

```

In [18]:

```
model_kernel3.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.5425 - sparse_categorical_accuracy: 0.8538 - 233ms/epoch - 745us/step
```

Out[18]:

```
[0.5425167679786682, 0.8537999987602234]
```

Model with random_normal initializer

In [19]:

```

model_kernel4 = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_normal'),
    tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_normal'),
    tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_normal')
])

```

In [20]:

```
model_kernel4.compile(optimizer='adam',  
                      loss='sparse_categorical_crossentropy',  
                      metrics=['sparse_categorical_accuracy'])
```

In [21]:

```
history_kernel4 = model_kernel4.fit(x_train, y_train,  
                                    batch_size=128,  
                                    epochs=100,  
                                    validation_data=(x_test, y_test),  
                                    verbose=2  
                                    )
```

```
Epoch 1/100  
469/469 - 1s - loss: 0.6770 - sparse_categorical_accuracy: 0.7573 - val_loss: 0.5413 - val_sparse_ca  
tegorical_accuracy: 0.8115 - 1s/epoch - 3ms/step  
Epoch 2/100  
469/469 - 1s - loss: 0.4790 - sparse_categorical_accuracy: 0.8289 - val_loss: 0.4722 - val_sparse_ca  
tegorical_accuracy: 0.8310 - 799ms/epoch - 2ms/step  
Epoch 3/100  
469/469 - 1s - loss: 0.4416 - sparse_categorical_accuracy: 0.8415 - val_loss: 0.4620 - val_sparse_ca  
tegorical_accuracy: 0.8347 - 841ms/epoch - 2ms/step  
Epoch 4/100  
469/469 - 1s - loss: 0.4218 - sparse_categorical_accuracy: 0.8473 - val_loss: 0.4568 - val_sparse_ca  
tegorical_accuracy: 0.8353 - 824ms/epoch - 2ms/step  
Epoch 5/100  
469/469 - 1s - loss: 0.4071 - sparse_categorical_accuracy: 0.8510 - val_loss: 0.4534 - val_sparse_ca  
tegorical_accuracy: 0.8391 - 756ms/epoch - 2ms/step  
Epoch 6/100  
469/469 - 1s - loss: 0.3950 - sparse_categorical_accuracy: 0.8558 - val_loss: 0.4365 - val_sparse_ca  
tegorical_accuracy: 0.8426 - 788ms/epoch - 2ms/step  
Epoch 7/100  
469/469 - 1s - loss: 0.3880 - sparse_categorical_accuracy: 0.8579 - val_loss: 0.4456 - val_sparse_ca  
tegorical_accuracy: 0.8454 - 800ms/epoch - 2ms/step  
Epoch 8/100  
469/469 - 1s - loss: 0.3813 - sparse_categorical_accuracy: 0.8590 - val_loss: 0.4275 - val_sparse_ca  
tegorical_accuracy: 0.8485 - 824ms/epoch - 2ms/step  
Epoch 9/100  
469/469 - 1s - loss: 0.3673 - sparse_categorical_accuracy: 0.8644 - val_loss: 0.4246 - val_sparse_ca  
tegorical_accuracy: 0.8517 - 796ms/epoch - 2ms/step  
Epoch 10/100  
469/469 - 1s - loss: 0.3639 - sparse_categorical_accuracy: 0.8660 - val_loss: 0.4441 - val_sparse_ca  
tegorical_accuracy: 0.8423 - 757ms/epoch - 2ms/step  
Epoch 11/100  
469/469 - 1s - loss: 0.3605 - sparse_categorical_accuracy: 0.8674 - val_loss: 0.4176 - val_sparse_ca  
tegorical_accuracy: 0.8479 - 755ms/epoch - 2ms/step  
Epoch 12/100  
469/469 - 1s - loss: 0.3547 - sparse_categorical_accuracy: 0.8697 - val_loss: 0.4194 - val_sparse_ca  
tegorical_accuracy: 0.8511 - 803ms/epoch - 2ms/step  
Epoch 13/100  
469/469 - 1s - loss: 0.3515 - sparse_categorical_accuracy: 0.8706 - val_loss: 0.4214 - val_sparse_ca  
tegorical_accuracy: 0.8510 - 799ms/epoch - 2ms/step  
Epoch 14/100  
469/469 - 1s - loss: 0.3476 - sparse_categorical_accuracy: 0.8726 - val_loss: 0.4280 - val_sparse_ca  
tegorical_accuracy: 0.8416 - 740ms/epoch - 2ms/step  
Epoch 15/100  
469/469 - 1s - loss: 0.3462 - sparse_categorical_accuracy: 0.8709 - val_loss: 0.4409 - val_sparse_ca  
tegorical_accuracy: 0.8399 - 777ms/epoch - 2ms/step  
Epoch 16/100  
469/469 - 1s - loss: 0.3397 - sparse_categorical_accuracy: 0.8737 - val_loss: 0.4092 - val_sparse_ca  
tegorical_accuracy: 0.8574 - 798ms/epoch - 2ms/step  
Epoch 17/100  
469/469 - 1s - loss: 0.3390 - sparse_categorical_accuracy: 0.8735 - val_loss: 0.4081 - val_sparse_ca  
tegorical_accuracy: 0.8530 - 837ms/epoch - 2ms/step  
Epoch 18/100  
469/469 - 1s - loss: 0.3340 - sparse_categorical_accuracy: 0.8757 - val_loss: 0.4217 - val_sparse_ca  
tegorical_accuracy: 0.8529 - 747ms/epoch - 2ms/step  
Epoch 19/100  
469/469 - 1s - loss: 0.3308 - sparse_categorical_accuracy: 0.8773 - val_loss: 0.4374 - val_sparse_ca  
tegorical_accuracy: 0.8424 - 736ms/epoch - 2ms/step  
Epoch 20/100  
469/469 - 1s - loss: 0.3285 - sparse_categorical_accuracy: 0.8782 - val_loss: 0.4303 - val_sparse_ca  
tegorical_accuracy: 0.8489 - 838ms/epoch - 2ms/step  
Epoch 21/100  
469/469 - 1s - loss: 0.3257 - sparse_categorical_accuracy: 0.8794 - val_loss: 0.4154 - val_sparse_ca  
tegorical_accuracy: 0.8562 - 758ms/epoch - 2ms/step  
Epoch 22/100  
469/469 - 1s - loss: 0.3238 - sparse_categorical_accuracy: 0.8795 - val_loss: 0.4241 - val_sparse_ca  
tegorical_accuracy: 0.8489 - 743ms/epoch - 2ms/step  
Epoch 23/100
```

469/469 - 1s - loss: 0.3272 - sparse_categorical_accuracy: 0.8784 - val_loss: 0.4144 - val_sparse_categorical_accuracy: 0.8521 - 788ms/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.3228 - sparse_categorical_accuracy: 0.8801 - val_loss: 0.4279 - val_sparse_categorical_accuracy: 0.8505 - 813ms/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.3220 - sparse_categorical_accuracy: 0.8799 - val_loss: 0.4749 - val_sparse_categorical_accuracy: 0.8322 - 736ms/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.3178 - sparse_categorical_accuracy: 0.8817 - val_loss: 0.4276 - val_sparse_categorical_accuracy: 0.8531 - 805ms/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.3182 - sparse_categorical_accuracy: 0.8828 - val_loss: 0.4148 - val_sparse_categorical_accuracy: 0.8579 - 792ms/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.3138 - sparse_categorical_accuracy: 0.8837 - val_loss: 0.4536 - val_sparse_categorical_accuracy: 0.8385 - 809ms/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.3163 - sparse_categorical_accuracy: 0.8814 - val_loss: 0.4311 - val_sparse_categorical_accuracy: 0.8511 - 775ms/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.3094 - sparse_categorical_accuracy: 0.8838 - val_loss: 0.4859 - val_sparse_categorical_accuracy: 0.8424 - 767ms/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.3160 - sparse_categorical_accuracy: 0.8816 - val_loss: 0.4561 - val_sparse_categorical_accuracy: 0.8458 - 750ms/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.3122 - sparse_categorical_accuracy: 0.8823 - val_loss: 0.4202 - val_sparse_categorical_accuracy: 0.8562 - 810ms/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.3102 - sparse_categorical_accuracy: 0.8841 - val_loss: 0.4257 - val_sparse_categorical_accuracy: 0.8558 - 801ms/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.3095 - sparse_categorical_accuracy: 0.8851 - val_loss: 0.4243 - val_sparse_categorical_accuracy: 0.8575 - 826ms/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.3028 - sparse_categorical_accuracy: 0.8863 - val_loss: 0.4354 - val_sparse_categorical_accuracy: 0.8532 - 743ms/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.3091 - sparse_categorical_accuracy: 0.8834 - val_loss: 0.4577 - val_sparse_categorical_accuracy: 0.8467 - 768ms/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.3054 - sparse_categorical_accuracy: 0.8848 - val_loss: 0.4376 - val_sparse_categorical_accuracy: 0.8597 - 732ms/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.3029 - sparse_categorical_accuracy: 0.8867 - val_loss: 0.4408 - val_sparse_categorical_accuracy: 0.8494 - 799ms/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.3044 - sparse_categorical_accuracy: 0.8856 - val_loss: 0.4521 - val_sparse_categorical_accuracy: 0.8509 - 769ms/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 0.3019 - sparse_categorical_accuracy: 0.8859 - val_loss: 0.4518 - val_sparse_categorical_accuracy: 0.8499 - 803ms/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.3005 - sparse_categorical_accuracy: 0.8858 - val_loss: 0.4406 - val_sparse_categorical_accuracy: 0.8515 - 748ms/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.3010 - sparse_categorical_accuracy: 0.8865 - val_loss: 0.4533 - val_sparse_categorical_accuracy: 0.8534 - 727ms/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.2996 - sparse_categorical_accuracy: 0.8882 - val_loss: 0.4801 - val_sparse_categorical_accuracy: 0.8404 - 776ms/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.2949 - sparse_categorical_accuracy: 0.8890 - val_loss: 0.4541 - val_sparse_categorical_accuracy: 0.8534 - 734ms/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.2997 - sparse_categorical_accuracy: 0.8870 - val_loss: 0.4512 - val_sparse_categorical_accuracy: 0.8511 - 802ms/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.2972 - sparse_categorical_accuracy: 0.8880 - val_loss: 0.4453 - val_sparse_categorical_accuracy: 0.8512 - 830ms/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.2978 - sparse_categorical_accuracy: 0.8882 - val_loss: 0.4362 - val_sparse_categorical_accuracy: 0.8567 - 737ms/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.2930 - sparse_categorical_accuracy: 0.8895 - val_loss: 0.4793 - val_sparse_categorical_accuracy: 0.8528 - 821ms/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.2946 - sparse_categorical_accuracy: 0.8895 - val_loss: 0.4477 - val_sparse_categorical_accuracy: 0.8556 - 724ms/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.2954 - sparse_categorical_accuracy: 0.8880 - val_loss: 0.4399 - val_sparse_categorical_accuracy: 0.8575 - 738ms/epoch - 2ms/step

Epoch 51/100
469/469 - 1s - loss: 0.2935 - sparse_categorical_accuracy: 0.8882 - val_loss: 0.4567 - val_sparse_categorical_accuracy: 0.8486 - 743ms/epoch - 2ms/step

Epoch 52/100
469/469 - 1s - loss: 0.2925 - sparse_categorical_accuracy: 0.8903 - val_loss: 0.4699 - val_sparse_categorical_accuracy: 0.8501 - 763ms/epoch - 2ms/step

Epoch 53/100
469/469 - 1s - loss: 0.2933 - sparse_categorical_accuracy: 0.8893 - val_loss: 0.4739 - val_sparse_categorical_accuracy: 0.8483 - 748ms/epoch - 2ms/step

Epoch 54/100
469/469 - 1s - loss: 0.2892 - sparse_categorical_accuracy: 0.8914 - val_loss: 0.4666 - val_sparse_categorical_accuracy: 0.8500 - 756ms/epoch - 2ms/step

Epoch 55/100
469/469 - 1s - loss: 0.2906 - sparse_categorical_accuracy: 0.8910 - val_loss: 0.4651 - val_sparse_categorical_accuracy: 0.8493 - 861ms/epoch - 2ms/step

Epoch 56/100
469/469 - 1s - loss: 0.2919 - sparse_categorical_accuracy: 0.8905 - val_loss: 0.4633 - val_sparse_categorical_accuracy: 0.8520 - 787ms/epoch - 2ms/step

Epoch 57/100
469/469 - 1s - loss: 0.2869 - sparse_categorical_accuracy: 0.8918 - val_loss: 0.4600 - val_sparse_categorical_accuracy: 0.8503 - 801ms/epoch - 2ms/step

Epoch 58/100
469/469 - 1s - loss: 0.2916 - sparse_categorical_accuracy: 0.8902 - val_loss: 0.4664 - val_sparse_categorical_accuracy: 0.8525 - 810ms/epoch - 2ms/step

Epoch 59/100
469/469 - 1s - loss: 0.2864 - sparse_categorical_accuracy: 0.8934 - val_loss: 0.4575 - val_sparse_categorical_accuracy: 0.8582 - 743ms/epoch - 2ms/step

Epoch 60/100
469/469 - 1s - loss: 0.2869 - sparse_categorical_accuracy: 0.8918 - val_loss: 0.4640 - val_sparse_categorical_accuracy: 0.8486 - 750ms/epoch - 2ms/step

Epoch 61/100
469/469 - 1s - loss: 0.2867 - sparse_categorical_accuracy: 0.8910 - val_loss: 0.4571 - val_sparse_categorical_accuracy: 0.8533 - 775ms/epoch - 2ms/step

Epoch 62/100
469/469 - 1s - loss: 0.2876 - sparse_categorical_accuracy: 0.8925 - val_loss: 0.4531 - val_sparse_categorical_accuracy: 0.8555 - 785ms/epoch - 2ms/step

Epoch 63/100
469/469 - 1s - loss: 0.2866 - sparse_categorical_accuracy: 0.8913 - val_loss: 0.4757 - val_sparse_categorical_accuracy: 0.8512 - 754ms/epoch - 2ms/step

Epoch 64/100
469/469 - 1s - loss: 0.2881 - sparse_categorical_accuracy: 0.8908 - val_loss: 0.4606 - val_sparse_categorical_accuracy: 0.8568 - 823ms/epoch - 2ms/step

Epoch 65/100
469/469 - 1s - loss: 0.2850 - sparse_categorical_accuracy: 0.8918 - val_loss: 0.4824 - val_sparse_categorical_accuracy: 0.8422 - 819ms/epoch - 2ms/step

Epoch 66/100
469/469 - 1s - loss: 0.2890 - sparse_categorical_accuracy: 0.8914 - val_loss: 0.4589 - val_sparse_categorical_accuracy: 0.8561 - 814ms/epoch - 2ms/step

Epoch 67/100
469/469 - 1s - loss: 0.2836 - sparse_categorical_accuracy: 0.8937 - val_loss: 0.4538 - val_sparse_categorical_accuracy: 0.8562 - 777ms/epoch - 2ms/step

Epoch 68/100
469/469 - 1s - loss: 0.2832 - sparse_categorical_accuracy: 0.8933 - val_loss: 0.4674 - val_sparse_categorical_accuracy: 0.8505 - 731ms/epoch - 2ms/step

Epoch 69/100
469/469 - 1s - loss: 0.2857 - sparse_categorical_accuracy: 0.8924 - val_loss: 0.4721 - val_sparse_categorical_accuracy: 0.8481 - 768ms/epoch - 2ms/step

Epoch 70/100
469/469 - 1s - loss: 0.2822 - sparse_categorical_accuracy: 0.8945 - val_loss: 0.4779 - val_sparse_categorical_accuracy: 0.8579 - 778ms/epoch - 2ms/step

Epoch 71/100
469/469 - 1s - loss: 0.2810 - sparse_categorical_accuracy: 0.8944 - val_loss: 0.4816 - val_sparse_categorical_accuracy: 0.8482 - 807ms/epoch - 2ms/step

Epoch 72/100
469/469 - 1s - loss: 0.2809 - sparse_categorical_accuracy: 0.8944 - val_loss: 0.4765 - val_sparse_categorical_accuracy: 0.8547 - 740ms/epoch - 2ms/step

Epoch 73/100
469/469 - 1s - loss: 0.2799 - sparse_categorical_accuracy: 0.8942 - val_loss: 0.4871 - val_sparse_categorical_accuracy: 0.8533 - 789ms/epoch - 2ms/step

Epoch 74/100
469/469 - 1s - loss: 0.2806 - sparse_categorical_accuracy: 0.8939 - val_loss: 0.4744 - val_sparse_categorical_accuracy: 0.8525 - 754ms/epoch - 2ms/step

Epoch 75/100
469/469 - 1s - loss: 0.2808 - sparse_categorical_accuracy: 0.8931 - val_loss: 0.4967 - val_sparse_categorical_accuracy: 0.8517 - 751ms/epoch - 2ms/step

Epoch 76/100
469/469 - 1s - loss: 0.2798 - sparse_categorical_accuracy: 0.8947 - val_loss: 0.4879 - val_sparse_categorical_accuracy: 0.8555 - 801ms/epoch - 2ms/step

Epoch 77/100
469/469 - 1s - loss: 0.2805 - sparse_categorical_accuracy: 0.8941 - val_loss: 0.4900 - val_sparse_categorical_accuracy: 0.8495 - 798ms/epoch - 2ms/step

Epoch 78/100
469/469 - 1s - loss: 0.2805 - sparse_categorical_accuracy: 0.8943 - val_loss: 0.4728 - val_sparse_categorical_accuracy: 0.8495 - 798ms/epoch - 2ms/step

```

tegorical_accuracy: 0.8528 - 749ms/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.2787 - sparse_categorical_accuracy: 0.8941 - val_loss: 0.4772 - val_sparse_ca
tegorical_accuracy: 0.8547 - 869ms/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.2783 - sparse_categorical_accuracy: 0.8954 - val_loss: 0.4722 - val_sparse_ca
tegorical_accuracy: 0.8530 - 737ms/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.2791 - sparse_categorical_accuracy: 0.8940 - val_loss: 0.5041 - val_sparse_ca
tegorical_accuracy: 0.8498 - 733ms/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.2787 - sparse_categorical_accuracy: 0.8936 - val_loss: 0.4860 - val_sparse_ca
tegorical_accuracy: 0.8524 - 850ms/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.2758 - sparse_categorical_accuracy: 0.8955 - val_loss: 0.4868 - val_sparse_ca
tegorical_accuracy: 0.8552 - 759ms/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.2804 - sparse_categorical_accuracy: 0.8945 - val_loss: 0.4827 - val_sparse_ca
tegorical_accuracy: 0.8520 - 730ms/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.2776 - sparse_categorical_accuracy: 0.8949 - val_loss: 0.4792 - val_sparse_ca
tegorical_accuracy: 0.8567 - 817ms/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.2791 - sparse_categorical_accuracy: 0.8958 - val_loss: 0.4894 - val_sparse_ca
tegorical_accuracy: 0.8461 - 790ms/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.2777 - sparse_categorical_accuracy: 0.8953 - val_loss: 0.5165 - val_sparse_ca
tegorical_accuracy: 0.8449 - 742ms/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.2785 - sparse_categorical_accuracy: 0.8954 - val_loss: 0.4993 - val_sparse_ca
tegorical_accuracy: 0.8424 - 811ms/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.2733 - sparse_categorical_accuracy: 0.8962 - val_loss: 0.4921 - val_sparse_ca
tegorical_accuracy: 0.8528 - 746ms/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.2719 - sparse_categorical_accuracy: 0.8961 - val_loss: 0.5067 - val_sparse_ca
tegorical_accuracy: 0.8493 - 692ms/epoch - 1ms/step
Epoch 91/100
469/469 - 1s - loss: 0.2752 - sparse_categorical_accuracy: 0.8951 - val_loss: 0.5076 - val_sparse_ca
tegorical_accuracy: 0.8504 - 730ms/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.2757 - sparse_categorical_accuracy: 0.8955 - val_loss: 0.5230 - val_sparse_ca
tegorical_accuracy: 0.8502 - 723ms/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.2746 - sparse_categorical_accuracy: 0.8967 - val_loss: 0.5191 - val_sparse_ca
tegorical_accuracy: 0.8445 - 799ms/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.2773 - sparse_categorical_accuracy: 0.8954 - val_loss: 0.5122 - val_sparse_ca
tegorical_accuracy: 0.8440 - 745ms/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.2761 - sparse_categorical_accuracy: 0.8958 - val_loss: 0.4927 - val_sparse_ca
tegorical_accuracy: 0.8516 - 792ms/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.2729 - sparse_categorical_accuracy: 0.8956 - val_loss: 0.5500 - val_sparse_ca
tegorical_accuracy: 0.8490 - 798ms/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.2701 - sparse_categorical_accuracy: 0.8979 - val_loss: 0.5028 - val_sparse_ca
tegorical_accuracy: 0.8443 - 723ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.2726 - sparse_categorical_accuracy: 0.8969 - val_loss: 0.5070 - val_sparse_ca
tegorical_accuracy: 0.8523 - 726ms/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.2702 - sparse_categorical_accuracy: 0.8982 - val_loss: 0.5198 - val_sparse_ca
tegorical_accuracy: 0.8464 - 705ms/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.2761 - sparse_categorical_accuracy: 0.8952 - val_loss: 0.5059 - val_sparse_ca
tegorical_accuracy: 0.8488 - 803ms/epoch - 2ms/step

```

In [22]:

```
model_kernel4.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.5059 - sparse_categorical_accuracy: 0.8488 - 243ms/epoch - 775us/step
```

Out[22]:

```
[0.5058592557907104, 0.848800003528595]
```

Model with he_uniform initializer

In [23]:

```
model_kernel5 = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(16, activation='relu', kernel_initializer='he_uniform'),
    tf.keras.layers.Dense(16, activation='relu', kernel_initializer='he_uniform'),
    tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='he_uniform')
])
```

In [24]:

```
model_kernel5.compile(optimizer='adam',
    loss='sparse_categorical_crossentropy',
    metrics=['sparse_categorical_accuracy']
)
```

In [25]:

```
history_kernel5 = model_kernel5.fit(x_train, y_train,
    batch_size=128,
    epochs=100,
    validation_data=(x_test, y_test),
    verbose=2
)
```

```
Epoch 1/100
469/469 - 1s - loss: 4.3028 - sparse_categorical_accuracy: 0.1338 - val_loss: 1.9935 - val_sparse_categorical_accuracy: 0.2132 - 1s/epoch - 2ms/step
Epoch 2/100
469/469 - 1s - loss: 1.8582 - sparse_categorical_accuracy: 0.2373 - val_loss: 1.7653 - val_sparse_categorical_accuracy: 0.2586 - 923ms/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 1.6922 - sparse_categorical_accuracy: 0.2953 - val_loss: 1.6745 - val_sparse_categorical_accuracy: 0.3275 - 802ms/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 1.5338 - sparse_categorical_accuracy: 0.3668 - val_loss: 1.4831 - val_sparse_categorical_accuracy: 0.3815 - 809ms/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 1.4217 - sparse_categorical_accuracy: 0.4042 - val_loss: 1.4057 - val_sparse_categorical_accuracy: 0.4175 - 743ms/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 1.3664 - sparse_categorical_accuracy: 0.4236 - val_loss: 1.3751 - val_sparse_categorical_accuracy: 0.4407 - 770ms/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 1.3282 - sparse_categorical_accuracy: 0.4361 - val_loss: 1.3397 - val_sparse_categorical_accuracy: 0.4344 - 749ms/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 1.2981 - sparse_categorical_accuracy: 0.4435 - val_loss: 1.2859 - val_sparse_categorical_accuracy: 0.4511 - 851ms/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 1.2711 - sparse_categorical_accuracy: 0.4480 - val_loss: 1.2767 - val_sparse_categorical_accuracy: 0.4563 - 725ms/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 1.2557 - sparse_categorical_accuracy: 0.4647 - val_loss: 1.2199 - val_sparse_categorical_accuracy: 0.5001 - 796ms/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 1.1642 - sparse_categorical_accuracy: 0.5080 - val_loss: 1.1576 - val_sparse_categorical_accuracy: 0.5226 - 766ms/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 1.1008 - sparse_categorical_accuracy: 0.5303 - val_loss: 1.0865 - val_sparse_categorical_accuracy: 0.5403 - 756ms/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 1.0890 - sparse_categorical_accuracy: 0.5296 - val_loss: 1.0629 - val_sparse_categorical_accuracy: 0.5366 - 759ms/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 1.0632 - sparse_categorical_accuracy: 0.5410 - val_loss: 1.1222 - val_sparse_categorical_accuracy: 0.5342 - 807ms/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 1.0664 - sparse_categorical_accuracy: 0.5423 - val_loss: 1.0605 - val_sparse_categorical_accuracy: 0.5469 - 750ms/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 1.0560 - sparse_categorical_accuracy: 0.5466 - val_loss: 1.0494 - val_sparse_categorical_accuracy: 0.5377 - 807ms/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 1.0416 - sparse_categorical_accuracy: 0.5474 - val_loss: 1.0686 - val_sparse_categorical_accuracy: 0.5533 - 796ms/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 1.0420 - sparse_categorical_accuracy: 0.5542 - val_loss: 1.0412 - val_sparse_categorical_accuracy: 0.5536 - 743ms/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 1.0501 - sparse_categorical_accuracy: 0.5529 - val_loss: 1.0495 - val_sparse_categorical_accuracy: 0.5455 - 830ms/epoch - 2ms/step
Epoch 20/100
```

469/469 - 1s - loss: 1.0352 - sparse_categorical_accuracy: 0.5549 - val_loss: 1.0286 - val_sparse_categorical_accuracy: 0.5633 - 788ms/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 1.0388 - sparse_categorical_accuracy: 0.5562 - val_loss: 1.0843 - val_sparse_categorical_accuracy: 0.5563 - 793ms/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 1.0376 - sparse_categorical_accuracy: 0.5567 - val_loss: 1.0179 - val_sparse_categorical_accuracy: 0.5698 - 716ms/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 1.0235 - sparse_categorical_accuracy: 0.5613 - val_loss: 1.0431 - val_sparse_categorical_accuracy: 0.5621 - 709ms/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 1.0185 - sparse_categorical_accuracy: 0.5603 - val_loss: 1.0206 - val_sparse_categorical_accuracy: 0.5569 - 787ms/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 1.0174 - sparse_categorical_accuracy: 0.5644 - val_loss: 1.0607 - val_sparse_categorical_accuracy: 0.5530 - 794ms/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 1.0376 - sparse_categorical_accuracy: 0.5539 - val_loss: 1.0652 - val_sparse_categorical_accuracy: 0.5410 - 715ms/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 1.0273 - sparse_categorical_accuracy: 0.5592 - val_loss: 1.0991 - val_sparse_categorical_accuracy: 0.5594 - 732ms/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 1.0202 - sparse_categorical_accuracy: 0.5653 - val_loss: 1.0162 - val_sparse_categorical_accuracy: 0.5575 - 731ms/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 1.0196 - sparse_categorical_accuracy: 0.5633 - val_loss: 1.0227 - val_sparse_categorical_accuracy: 0.5598 - 799ms/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 1.0284 - sparse_categorical_accuracy: 0.5594 - val_loss: 1.0375 - val_sparse_categorical_accuracy: 0.5769 - 794ms/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 1.0178 - sparse_categorical_accuracy: 0.5637 - val_loss: 1.0505 - val_sparse_categorical_accuracy: 0.5684 - 745ms/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 1.0125 - sparse_categorical_accuracy: 0.5659 - val_loss: 1.0368 - val_sparse_categorical_accuracy: 0.5710 - 796ms/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 1.0212 - sparse_categorical_accuracy: 0.5652 - val_loss: 1.0225 - val_sparse_categorical_accuracy: 0.5751 - 804ms/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 1.0158 - sparse_categorical_accuracy: 0.5694 - val_loss: 1.0475 - val_sparse_categorical_accuracy: 0.5588 - 808ms/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 1.0153 - sparse_categorical_accuracy: 0.5688 - val_loss: 1.0570 - val_sparse_categorical_accuracy: 0.5548 - 814ms/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 1.0269 - sparse_categorical_accuracy: 0.5617 - val_loss: 1.0293 - val_sparse_categorical_accuracy: 0.5700 - 818ms/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 1.0228 - sparse_categorical_accuracy: 0.5612 - val_loss: 1.0106 - val_sparse_categorical_accuracy: 0.5696 - 714ms/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 1.0153 - sparse_categorical_accuracy: 0.5677 - val_loss: 1.0508 - val_sparse_categorical_accuracy: 0.5644 - 767ms/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 1.0266 - sparse_categorical_accuracy: 0.5617 - val_loss: 1.0308 - val_sparse_categorical_accuracy: 0.5449 - 738ms/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 1.0160 - sparse_categorical_accuracy: 0.5620 - val_loss: 1.0298 - val_sparse_categorical_accuracy: 0.5514 - 798ms/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 1.0242 - sparse_categorical_accuracy: 0.5590 - val_loss: 1.0278 - val_sparse_categorical_accuracy: 0.5627 - 837ms/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 1.0179 - sparse_categorical_accuracy: 0.5663 - val_loss: 1.0398 - val_sparse_categorical_accuracy: 0.5251 - 810ms/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.9996 - sparse_categorical_accuracy: 0.5723 - val_loss: 1.0206 - val_sparse_categorical_accuracy: 0.5700 - 746ms/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 1.0117 - sparse_categorical_accuracy: 0.5630 - val_loss: 1.0453 - val_sparse_categorical_accuracy: 0.5434 - 723ms/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 1.0189 - sparse_categorical_accuracy: 0.5641 - val_loss: 1.0545 - val_sparse_categorical_accuracy: 0.5204 - 786ms/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 1.0200 - sparse_categorical_accuracy: 0.5650 - val_loss: 1.0338 - val_sparse_categorical_accuracy: 0.5303 - 762ms/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 1.0262 - sparse_categorical_accuracy: 0.5586 - val_loss: 1.0230 - val_sparse_categorical_accuracy: 0.5668 - 743ms/epoch - 2ms/step

Epoch 48/100
469/469 - 1s - loss: 1.0054 - sparse_categorical_accuracy: 0.5709 - val_loss: 1.0391 - val_sparse_categorical_accuracy: 0.5481 - 821ms/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 1.0000 - sparse_categorical_accuracy: 0.5736 - val_loss: 1.0501 - val_sparse_categorical_accuracy: 0.5584 - 740ms/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 1.0251 - sparse_categorical_accuracy: 0.5602 - val_loss: 1.0486 - val_sparse_categorical_accuracy: 0.5659 - 773ms/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 1.0102 - sparse_categorical_accuracy: 0.5692 - val_loss: 1.0252 - val_sparse_categorical_accuracy: 0.5743 - 751ms/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 1.0096 - sparse_categorical_accuracy: 0.5726 - val_loss: 1.0305 - val_sparse_categorical_accuracy: 0.5664 - 824ms/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 1.0335 - sparse_categorical_accuracy: 0.5588 - val_loss: 1.0249 - val_sparse_categorical_accuracy: 0.5604 - 864ms/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 1.0223 - sparse_categorical_accuracy: 0.5627 - val_loss: 1.0153 - val_sparse_categorical_accuracy: 0.5660 - 764ms/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 1.0165 - sparse_categorical_accuracy: 0.5691 - val_loss: 1.0165 - val_sparse_categorical_accuracy: 0.5750 - 736ms/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.9767 - sparse_categorical_accuracy: 0.5814 - val_loss: 1.1381 - val_sparse_categorical_accuracy: 0.5399 - 704ms/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 1.0216 - sparse_categorical_accuracy: 0.5660 - val_loss: 1.0502 - val_sparse_categorical_accuracy: 0.5553 - 742ms/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 1.0202 - sparse_categorical_accuracy: 0.5638 - val_loss: 0.9943 - val_sparse_categorical_accuracy: 0.5789 - 751ms/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 1.0064 - sparse_categorical_accuracy: 0.5683 - val_loss: 1.0008 - val_sparse_categorical_accuracy: 0.5753 - 753ms/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 1.0192 - sparse_categorical_accuracy: 0.5635 - val_loss: 1.0340 - val_sparse_categorical_accuracy: 0.5432 - 715ms/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 1.0047 - sparse_categorical_accuracy: 0.5678 - val_loss: 1.0038 - val_sparse_categorical_accuracy: 0.5784 - 828ms/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 1.0103 - sparse_categorical_accuracy: 0.5718 - val_loss: 0.9968 - val_sparse_categorical_accuracy: 0.5767 - 838ms/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 1.0112 - sparse_categorical_accuracy: 0.5728 - val_loss: 1.0606 - val_sparse_categorical_accuracy: 0.5615 - 796ms/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 1.0090 - sparse_categorical_accuracy: 0.5680 - val_loss: 1.0097 - val_sparse_categorical_accuracy: 0.5643 - 698ms/epoch - 1ms/step
Epoch 65/100
469/469 - 1s - loss: 1.0479 - sparse_categorical_accuracy: 0.5544 - val_loss: 0.9966 - val_sparse_categorical_accuracy: 0.5865 - 732ms/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.9961 - sparse_categorical_accuracy: 0.5705 - val_loss: 1.0188 - val_sparse_categorical_accuracy: 0.5770 - 783ms/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.9854 - sparse_categorical_accuracy: 0.5768 - val_loss: 1.1034 - val_sparse_categorical_accuracy: 0.5300 - 730ms/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 1.0140 - sparse_categorical_accuracy: 0.5661 - val_loss: 1.0095 - val_sparse_categorical_accuracy: 0.5755 - 721ms/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 1.0052 - sparse_categorical_accuracy: 0.5702 - val_loss: 1.0337 - val_sparse_categorical_accuracy: 0.5482 - 788ms/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 1.0029 - sparse_categorical_accuracy: 0.5682 - val_loss: 1.0287 - val_sparse_categorical_accuracy: 0.5336 - 808ms/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.9975 - sparse_categorical_accuracy: 0.5721 - val_loss: 0.9921 - val_sparse_categorical_accuracy: 0.5779 - 807ms/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 1.0533 - sparse_categorical_accuracy: 0.5596 - val_loss: 1.1599 - val_sparse_categorical_accuracy: 0.4969 - 719ms/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.9884 - sparse_categorical_accuracy: 0.5746 - val_loss: 1.0159 - val_sparse_categorical_accuracy: 0.5727 - 696ms/epoch - 1ms/step
Epoch 74/100
469/469 - 1s - loss: 0.9939 - sparse_categorical_accuracy: 0.5768 - val_loss: 1.0496 - val_sparse_categorical_accuracy: 0.5518 - 808ms/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 1.0293 - sparse_categorical_accuracy: 0.5621 - val_loss: 1.0001 - val_sparse_categorical_accuracy: 0.5621 - 808ms/epoch - 2ms/step

tegorical_accuracy: 0.5758 - 800ms/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.9927 - sparse_categorical_accuracy: 0.5792 - val_loss: 1.0346 - val_sparse_categorical_accuracy: 0.5683 - 769ms/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 1.0129 - sparse_categorical_accuracy: 0.5725 - val_loss: 1.0315 - val_sparse_categorical_accuracy: 0.6015 - 712ms/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.9935 - sparse_categorical_accuracy: 0.5860 - val_loss: 0.9963 - val_sparse_categorical_accuracy: 0.5780 - 744ms/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.9742 - sparse_categorical_accuracy: 0.5816 - val_loss: 1.0067 - val_sparse_categorical_accuracy: 0.5845 - 738ms/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 1.0081 - sparse_categorical_accuracy: 0.5698 - val_loss: 1.0659 - val_sparse_categorical_accuracy: 0.5463 - 835ms/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.9894 - sparse_categorical_accuracy: 0.5729 - val_loss: 1.0135 - val_sparse_categorical_accuracy: 0.5709 - 722ms/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.9972 - sparse_categorical_accuracy: 0.5735 - val_loss: 1.0118 - val_sparse_categorical_accuracy: 0.5759 - 773ms/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.9816 - sparse_categorical_accuracy: 0.5815 - val_loss: 1.0056 - val_sparse_categorical_accuracy: 0.5813 - 719ms/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.9902 - sparse_categorical_accuracy: 0.5798 - val_loss: 1.0448 - val_sparse_categorical_accuracy: 0.5499 - 796ms/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 1.0055 - sparse_categorical_accuracy: 0.5666 - val_loss: 0.9999 - val_sparse_categorical_accuracy: 0.5677 - 743ms/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.9931 - sparse_categorical_accuracy: 0.5748 - val_loss: 1.0377 - val_sparse_categorical_accuracy: 0.5672 - 685ms/epoch - 1ms/step
Epoch 87/100
469/469 - 1s - loss: 1.0220 - sparse_categorical_accuracy: 0.5688 - val_loss: 1.0113 - val_sparse_categorical_accuracy: 0.5671 - 740ms/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.9835 - sparse_categorical_accuracy: 0.5717 - val_loss: 0.9885 - val_sparse_categorical_accuracy: 0.5808 - 784ms/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.9984 - sparse_categorical_accuracy: 0.5808 - val_loss: 0.9955 - val_sparse_categorical_accuracy: 0.6021 - 801ms/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.9851 - sparse_categorical_accuracy: 0.5859 - val_loss: 1.0243 - val_sparse_categorical_accuracy: 0.5732 - 754ms/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.9807 - sparse_categorical_accuracy: 0.5924 - val_loss: 0.9729 - val_sparse_categorical_accuracy: 0.6127 - 830ms/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.9700 - sparse_categorical_accuracy: 0.5980 - val_loss: 0.9907 - val_sparse_categorical_accuracy: 0.5885 - 948ms/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.9997 - sparse_categorical_accuracy: 0.5931 - val_loss: 0.9592 - val_sparse_categorical_accuracy: 0.6223 - 916ms/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.9747 - sparse_categorical_accuracy: 0.6015 - val_loss: 0.9960 - val_sparse_categorical_accuracy: 0.6039 - 961ms/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.9491 - sparse_categorical_accuracy: 0.6056 - val_loss: 0.9710 - val_sparse_categorical_accuracy: 0.5954 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.9596 - sparse_categorical_accuracy: 0.5993 - val_loss: 0.9732 - val_sparse_categorical_accuracy: 0.6186 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.9642 - sparse_categorical_accuracy: 0.5990 - val_loss: 1.0066 - val_sparse_categorical_accuracy: 0.5713 - 781ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.9438 - sparse_categorical_accuracy: 0.6099 - val_loss: 0.9845 - val_sparse_categorical_accuracy: 0.5929 - 712ms/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.9532 - sparse_categorical_accuracy: 0.6043 - val_loss: 0.9661 - val_sparse_categorical_accuracy: 0.6184 - 789ms/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.9404 - sparse_categorical_accuracy: 0.6143 - val_loss: 0.9433 - val_sparse_categorical_accuracy: 0.6244 - 841ms/epoch - 2ms/step

In [26]:

```
model_kernel5.evaluate(x_test, y_test, verbose=2)
```

313/313 - 0s - loss: 0.9433 - sparse_categorical_accuracy: 0.6244 - 244ms/epoch - 779us/step

Out[26]:

```
[0.9432547688484192, 0.6244000196456909]
```

Model with he_normal initializer

In [27]:

```
model_kernel6 = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(16, activation='relu', kernel_initializer='he_normal'),
    tf.keras.layers.Dense(16, activation='relu', kernel_initializer='he_normal'),
    tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='he_normal')
])
```

In [28]:

```
model_kernel6.compile(optimizer='adam',
                      loss='sparse_categorical_crossentropy',
                      metrics=['sparse_categorical_accuracy'])
```

In [29]:

```
history_kernel6 = model_kernel6.fit(x_train, y_train,
                                    batch_size=128,
                                    epochs=100,
                                    validation_data=(x_test, y_test),
                                    verbose=2)
```

Epoch 1/100

469/469 - 1s - loss: 4.0071 - sparse_categorical_accuracy: 0.2628 - val_loss: 1.8795 - val_sparse_categorical_accuracy: 0.3103 - 1s/epoch - 3ms/step

Epoch 2/100

469/469 - 1s - loss: 1.7693 - sparse_categorical_accuracy: 0.3353 - val_loss: 1.6998 - val_sparse_categorical_accuracy: 0.3646 - 760ms/epoch - 2ms/step

Epoch 3/100

469/469 - 1s - loss: 1.5803 - sparse_categorical_accuracy: 0.4108 - val_loss: 1.4984 - val_sparse_categorical_accuracy: 0.4369 - 819ms/epoch - 2ms/step

Epoch 4/100

469/469 - 1s - loss: 1.4339 - sparse_categorical_accuracy: 0.4572 - val_loss: 1.4193 - val_sparse_categorical_accuracy: 0.4516 - 805ms/epoch - 2ms/step

Epoch 5/100

469/469 - 1s - loss: 1.2742 - sparse_categorical_accuracy: 0.5191 - val_loss: 1.3070 - val_sparse_categorical_accuracy: 0.5344 - 779ms/epoch - 2ms/step

Epoch 6/100

469/469 - 1s - loss: 1.1169 - sparse_categorical_accuracy: 0.5733 - val_loss: 1.1240 - val_sparse_categorical_accuracy: 0.5739 - 827ms/epoch - 2ms/step

Epoch 7/100

469/469 - 1s - loss: 1.0369 - sparse_categorical_accuracy: 0.5963 - val_loss: 1.1252 - val_sparse_categorical_accuracy: 0.5672 - 815ms/epoch - 2ms/step

Epoch 8/100

469/469 - 1s - loss: 0.9821 - sparse_categorical_accuracy: 0.6102 - val_loss: 0.9930 - val_sparse_categorical_accuracy: 0.6267 - 775ms/epoch - 2ms/step

Epoch 9/100

469/469 - 1s - loss: 0.9599 - sparse_categorical_accuracy: 0.6132 - val_loss: 0.9893 - val_sparse_categorical_accuracy: 0.6240 - 797ms/epoch - 2ms/step

Epoch 10/100

469/469 - 1s - loss: 0.9185 - sparse_categorical_accuracy: 0.6238 - val_loss: 1.0140 - val_sparse_categorical_accuracy: 0.6074 - 761ms/epoch - 2ms/step

Epoch 11/100

469/469 - 1s - loss: 0.8796 - sparse_categorical_accuracy: 0.6373 - val_loss: 0.9799 - val_sparse_categorical_accuracy: 0.6169 - 717ms/epoch - 2ms/step

Epoch 12/100

469/469 - 1s - loss: 0.8570 - sparse_categorical_accuracy: 0.6433 - val_loss: 0.8950 - val_sparse_categorical_accuracy: 0.6612 - 801ms/epoch - 2ms/step

Epoch 13/100

469/469 - 1s - loss: 0.8194 - sparse_categorical_accuracy: 0.6588 - val_loss: 0.8600 - val_sparse_categorical_accuracy: 0.6827 - 1s/epoch - 2ms/step

Epoch 14/100

469/469 - 1s - loss: 0.7702 - sparse_categorical_accuracy: 0.6929 - val_loss: 0.7530 - val_sparse_categorical_accuracy: 0.7136 - 1s/epoch - 2ms/step

Epoch 15/100

469/469 - 1s - loss: 0.6961 - sparse_categorical_accuracy: 0.7250 - val_loss: 0.7281 - val_sparse_categorical_accuracy: 0.7263 - 991ms/epoch - 2ms/step

Epoch 16/100
469/469 - 1s - loss: 0.6643 - sparse_categorical_accuracy: 0.7356 - val_loss: 0.8201 - val_sparse_categorical_accuracy: 0.6908 - 943ms/epoch - 2ms/step

Epoch 17/100
469/469 - 1s - loss: 0.6462 - sparse_categorical_accuracy: 0.7392 - val_loss: 0.7601 - val_sparse_categorical_accuracy: 0.6999 - 1s/epoch - 2ms/step

Epoch 18/100
469/469 - 1s - loss: 0.6385 - sparse_categorical_accuracy: 0.7421 - val_loss: 0.6831 - val_sparse_categorical_accuracy: 0.7329 - 1s/epoch - 2ms/step

Epoch 19/100
469/469 - 1s - loss: 0.6277 - sparse_categorical_accuracy: 0.7459 - val_loss: 0.6759 - val_sparse_categorical_accuracy: 0.7348 - 785ms/epoch - 2ms/step

Epoch 20/100
469/469 - 1s - loss: 0.6091 - sparse_categorical_accuracy: 0.7499 - val_loss: 0.6468 - val_sparse_categorical_accuracy: 0.7430 - 795ms/epoch - 2ms/step

Epoch 21/100
469/469 - 1s - loss: 0.6130 - sparse_categorical_accuracy: 0.7496 - val_loss: 0.6767 - val_sparse_categorical_accuracy: 0.7335 - 703ms/epoch - 1ms/step

Epoch 22/100
469/469 - 1s - loss: 0.5992 - sparse_categorical_accuracy: 0.7543 - val_loss: 0.6986 - val_sparse_categorical_accuracy: 0.7159 - 722ms/epoch - 2ms/step

Epoch 23/100
469/469 - 1s - loss: 0.5909 - sparse_categorical_accuracy: 0.7548 - val_loss: 0.6478 - val_sparse_categorical_accuracy: 0.7452 - 809ms/epoch - 2ms/step

Epoch 24/100
469/469 - 1s - loss: 0.5802 - sparse_categorical_accuracy: 0.7594 - val_loss: 0.6880 - val_sparse_categorical_accuracy: 0.7291 - 784ms/epoch - 2ms/step

Epoch 25/100
469/469 - 1s - loss: 0.5830 - sparse_categorical_accuracy: 0.7599 - val_loss: 0.6311 - val_sparse_categorical_accuracy: 0.7471 - 847ms/epoch - 2ms/step

Epoch 26/100
469/469 - 1s - loss: 0.5697 - sparse_categorical_accuracy: 0.7630 - val_loss: 0.6414 - val_sparse_categorical_accuracy: 0.7526 - 755ms/epoch - 2ms/step

Epoch 27/100
469/469 - 1s - loss: 0.5682 - sparse_categorical_accuracy: 0.7655 - val_loss: 0.6373 - val_sparse_categorical_accuracy: 0.7451 - 845ms/epoch - 2ms/step

Epoch 28/100
469/469 - 1s - loss: 0.5597 - sparse_categorical_accuracy: 0.7681 - val_loss: 0.6192 - val_sparse_categorical_accuracy: 0.7521 - 808ms/epoch - 2ms/step

Epoch 29/100
469/469 - 1s - loss: 0.5633 - sparse_categorical_accuracy: 0.7665 - val_loss: 0.6168 - val_sparse_categorical_accuracy: 0.7576 - 809ms/epoch - 2ms/step

Epoch 30/100
469/469 - 1s - loss: 0.5589 - sparse_categorical_accuracy: 0.7677 - val_loss: 0.6294 - val_sparse_categorical_accuracy: 0.7499 - 820ms/epoch - 2ms/step

Epoch 31/100
469/469 - 1s - loss: 0.5533 - sparse_categorical_accuracy: 0.7700 - val_loss: 0.6096 - val_sparse_categorical_accuracy: 0.7559 - 809ms/epoch - 2ms/step

Epoch 32/100
469/469 - 1s - loss: 0.5488 - sparse_categorical_accuracy: 0.7709 - val_loss: 0.6117 - val_sparse_categorical_accuracy: 0.7580 - 926ms/epoch - 2ms/step

Epoch 33/100
469/469 - 1s - loss: 0.5516 - sparse_categorical_accuracy: 0.7695 - val_loss: 0.6242 - val_sparse_categorical_accuracy: 0.7542 - 1s/epoch - 2ms/step

Epoch 34/100
469/469 - 1s - loss: 0.5481 - sparse_categorical_accuracy: 0.7708 - val_loss: 0.6518 - val_sparse_categorical_accuracy: 0.7560 - 773ms/epoch - 2ms/step

Epoch 35/100
469/469 - 1s - loss: 0.5435 - sparse_categorical_accuracy: 0.7734 - val_loss: 0.6309 - val_sparse_categorical_accuracy: 0.7564 - 779ms/epoch - 2ms/step

Epoch 36/100
469/469 - 1s - loss: 0.5427 - sparse_categorical_accuracy: 0.7744 - val_loss: 0.6477 - val_sparse_categorical_accuracy: 0.7541 - 954ms/epoch - 2ms/step

Epoch 37/100
469/469 - 1s - loss: 0.5407 - sparse_categorical_accuracy: 0.7735 - val_loss: 0.6132 - val_sparse_categorical_accuracy: 0.7602 - 1s/epoch - 2ms/step

Epoch 38/100
469/469 - 1s - loss: 0.5371 - sparse_categorical_accuracy: 0.7755 - val_loss: 0.6034 - val_sparse_categorical_accuracy: 0.7610 - 727ms/epoch - 2ms/step

Epoch 39/100
469/469 - 1s - loss: 0.5314 - sparse_categorical_accuracy: 0.7782 - val_loss: 0.6143 - val_sparse_categorical_accuracy: 0.7610 - 799ms/epoch - 2ms/step

Epoch 40/100
469/469 - 1s - loss: 0.5339 - sparse_categorical_accuracy: 0.7777 - val_loss: 0.6128 - val_sparse_categorical_accuracy: 0.7566 - 836ms/epoch - 2ms/step

Epoch 41/100
469/469 - 1s - loss: 0.5327 - sparse_categorical_accuracy: 0.7794 - val_loss: 0.6317 - val_sparse_categorical_accuracy: 0.7638 - 809ms/epoch - 2ms/step

Epoch 42/100
469/469 - 1s - loss: 0.5248 - sparse_categorical_accuracy: 0.7826 - val_loss: 0.6052 - val_sparse_categorical_accuracy: 0.7726 - 790ms/epoch - 2ms/step

Epoch 43/100
469/469 - 1s - loss: 0.5213 - sparse_categorical_accuracy: 0.7874 - val_loss: 0.6006 - val_sparse_categorical_accuracy: 0.7874 - 790ms/epoch - 2ms/step

tegorical_accuracy: 0.7702 - 996ms/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.5215 - sparse_categorical_accuracy: 0.7875 - val_loss: 0.6035 - val_sparse_ca
tegorical_accuracy: 0.7675 - 800ms/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.5154 - sparse_categorical_accuracy: 0.7897 - val_loss: 0.6207 - val_sparse_ca
tegorical_accuracy: 0.7760 - 864ms/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.5163 - sparse_categorical_accuracy: 0.7918 - val_loss: 0.6219 - val_sparse_ca
tegorical_accuracy: 0.7710 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.5141 - sparse_categorical_accuracy: 0.7911 - val_loss: 0.6452 - val_sparse_ca
tegorical_accuracy: 0.7627 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.5062 - sparse_categorical_accuracy: 0.7947 - val_loss: 0.6003 - val_sparse_ca
tegorical_accuracy: 0.7754 - 782ms/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.5066 - sparse_categorical_accuracy: 0.7971 - val_loss: 0.6280 - val_sparse_ca
tegorical_accuracy: 0.7861 - 811ms/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.5045 - sparse_categorical_accuracy: 0.8010 - val_loss: 0.6105 - val_sparse_ca
tegorical_accuracy: 0.7849 - 991ms/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.4978 - sparse_categorical_accuracy: 0.8082 - val_loss: 0.6185 - val_sparse_ca
tegorical_accuracy: 0.7914 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.4890 - sparse_categorical_accuracy: 0.8126 - val_loss: 0.6228 - val_sparse_ca
tegorical_accuracy: 0.7779 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.4874 - sparse_categorical_accuracy: 0.8143 - val_loss: 0.6071 - val_sparse_ca
tegorical_accuracy: 0.7864 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.4806 - sparse_categorical_accuracy: 0.8176 - val_loss: 0.5789 - val_sparse_ca
tegorical_accuracy: 0.7958 - 708ms/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.4739 - sparse_categorical_accuracy: 0.8216 - val_loss: 0.5681 - val_sparse_ca
tegorical_accuracy: 0.8031 - 749ms/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.4647 - sparse_categorical_accuracy: 0.8280 - val_loss: 0.5818 - val_sparse_ca
tegorical_accuracy: 0.8064 - 684ms/epoch - 1ms/step
Epoch 57/100
469/469 - 1s - loss: 0.4621 - sparse_categorical_accuracy: 0.8329 - val_loss: 0.5840 - val_sparse_ca
tegorical_accuracy: 0.8068 - 815ms/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.4613 - sparse_categorical_accuracy: 0.8371 - val_loss: 0.5767 - val_sparse_ca
tegorical_accuracy: 0.8068 - 742ms/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.4486 - sparse_categorical_accuracy: 0.8421 - val_loss: 0.5689 - val_sparse_ca
tegorical_accuracy: 0.8160 - 774ms/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.4423 - sparse_categorical_accuracy: 0.8444 - val_loss: 0.5457 - val_sparse_ca
tegorical_accuracy: 0.8217 - 752ms/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.4330 - sparse_categorical_accuracy: 0.8491 - val_loss: 0.5633 - val_sparse_ca
tegorical_accuracy: 0.8221 - 793ms/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.4290 - sparse_categorical_accuracy: 0.8500 - val_loss: 0.5758 - val_sparse_ca
tegorical_accuracy: 0.8079 - 724ms/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 0.4281 - sparse_categorical_accuracy: 0.8502 - val_loss: 0.5448 - val_sparse_ca
tegorical_accuracy: 0.8195 - 803ms/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.4231 - sparse_categorical_accuracy: 0.8503 - val_loss: 0.5402 - val_sparse_ca
tegorical_accuracy: 0.8239 - 722ms/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.4163 - sparse_categorical_accuracy: 0.8539 - val_loss: 0.5597 - val_sparse_ca
tegorical_accuracy: 0.8214 - 720ms/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.4157 - sparse_categorical_accuracy: 0.8547 - val_loss: 0.5268 - val_sparse_ca
tegorical_accuracy: 0.8221 - 814ms/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.4138 - sparse_categorical_accuracy: 0.8541 - val_loss: 0.5505 - val_sparse_ca
tegorical_accuracy: 0.8225 - 744ms/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 0.4119 - sparse_categorical_accuracy: 0.8545 - val_loss: 0.5652 - val_sparse_ca
tegorical_accuracy: 0.8208 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 0.4037 - sparse_categorical_accuracy: 0.8581 - val_loss: 0.5395 - val_sparse_ca
tegorical_accuracy: 0.8218 - 940ms/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.4017 - sparse_categorical_accuracy: 0.8579 - val_loss: 0.5488 - val_sparse_ca
tegorical_accuracy: 0.8217 - 720ms/epoch - 2ms/step
Epoch 71/100

469/469 - 1s - loss: 0.4026 - sparse_categorical_accuracy: 0.8565 - val_loss: 0.5280 - val_sparse_categorical_accuracy: 0.8309 - 844ms/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.3970 - sparse_categorical_accuracy: 0.8599 - val_loss: 0.5396 - val_sparse_categorical_accuracy: 0.8257 - 728ms/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.3950 - sparse_categorical_accuracy: 0.8594 - val_loss: 0.5346 - val_sparse_categorical_accuracy: 0.8288 - 737ms/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.3962 - sparse_categorical_accuracy: 0.8596 - val_loss: 0.5731 - val_sparse_categorical_accuracy: 0.8223 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.3995 - sparse_categorical_accuracy: 0.8584 - val_loss: 0.5427 - val_sparse_categorical_accuracy: 0.8214 - 819ms/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.3917 - sparse_categorical_accuracy: 0.8620 - val_loss: 0.5318 - val_sparse_categorical_accuracy: 0.8349 - 780ms/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.3932 - sparse_categorical_accuracy: 0.8608 - val_loss: 0.5286 - val_sparse_categorical_accuracy: 0.8242 - 790ms/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.3872 - sparse_categorical_accuracy: 0.8626 - val_loss: 0.5192 - val_sparse_categorical_accuracy: 0.8361 - 817ms/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.3886 - sparse_categorical_accuracy: 0.8619 - val_loss: 0.5449 - val_sparse_categorical_accuracy: 0.8295 - 796ms/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.3848 - sparse_categorical_accuracy: 0.8633 - val_loss: 0.5459 - val_sparse_categorical_accuracy: 0.8276 - 743ms/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.3898 - sparse_categorical_accuracy: 0.8611 - val_loss: 0.5579 - val_sparse_categorical_accuracy: 0.8216 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.3869 - sparse_categorical_accuracy: 0.8637 - val_loss: 0.5417 - val_sparse_categorical_accuracy: 0.8294 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.3834 - sparse_categorical_accuracy: 0.8636 - val_loss: 0.5563 - val_sparse_categorical_accuracy: 0.8291 - 931ms/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.3901 - sparse_categorical_accuracy: 0.8616 - val_loss: 0.5751 - val_sparse_categorical_accuracy: 0.8333 - 796ms/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.3807 - sparse_categorical_accuracy: 0.8649 - val_loss: 0.5916 - val_sparse_categorical_accuracy: 0.8183 - 781ms/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.3772 - sparse_categorical_accuracy: 0.8655 - val_loss: 0.5204 - val_sparse_categorical_accuracy: 0.8320 - 725ms/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.3760 - sparse_categorical_accuracy: 0.8651 - val_loss: 0.5381 - val_sparse_categorical_accuracy: 0.8322 - 807ms/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.3766 - sparse_categorical_accuracy: 0.8662 - val_loss: 0.5386 - val_sparse_categorical_accuracy: 0.8293 - 812ms/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.3771 - sparse_categorical_accuracy: 0.8655 - val_loss: 0.5840 - val_sparse_categorical_accuracy: 0.8022 - 814ms/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.3753 - sparse_categorical_accuracy: 0.8651 - val_loss: 0.5256 - val_sparse_categorical_accuracy: 0.8341 - 756ms/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.3755 - sparse_categorical_accuracy: 0.8656 - val_loss: 0.5441 - val_sparse_categorical_accuracy: 0.8340 - 788ms/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.3769 - sparse_categorical_accuracy: 0.8654 - val_loss: 0.5365 - val_sparse_categorical_accuracy: 0.8316 - 878ms/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.3726 - sparse_categorical_accuracy: 0.8665 - val_loss: 0.5522 - val_sparse_categorical_accuracy: 0.8256 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.3693 - sparse_categorical_accuracy: 0.8683 - val_loss: 0.5359 - val_sparse_categorical_accuracy: 0.8359 - 903ms/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.3712 - sparse_categorical_accuracy: 0.8677 - val_loss: 0.5790 - val_sparse_categorical_accuracy: 0.8255 - 861ms/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.3734 - sparse_categorical_accuracy: 0.8668 - val_loss: 0.5233 - val_sparse_categorical_accuracy: 0.8316 - 949ms/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.3735 - sparse_categorical_accuracy: 0.8673 - val_loss: 0.5366 - val_sparse_categorical_accuracy: 0.8262 - 925ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.3675 - sparse_categorical_accuracy: 0.8680 - val_loss: 0.5134 - val_sparse_categorical_accuracy: 0.8357 - 815ms/epoch - 2ms/step

```
Epoch 99/100
469/469 - 1s - loss: 0.3632 - sparse_categorical_accuracy: 0.8699 - val_loss: 0.5389 - val_sparse_categorical_accuracy: 0.8316 - 732ms/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.3629 - sparse_categorical_accuracy: 0.8697 - val_loss: 0.5581 - val_sparse_categorical_accuracy: 0.8241 - 816ms/epoch - 2ms/step
```

In [30]:

```
model_kernel6.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.5581 - sparse_categorical_accuracy: 0.8241 - 241ms/epoch - 771us/step
```

Out[30]:

```
[0.5580878853797913, 0.8241000175476074]
```

Try different bias initializers

Model with `glorot_normal` bias initializer

In [31]:

```
model_bias1 = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(16, activation='relu', bias_initializer='glorot_normal'),
    tf.keras.layers.Dense(16, activation='relu', bias_initializer='glorot_normal'),
    tf.keras.layers.Dense(10, activation='softmax', bias_initializer='glorot_normal')
])
```

In [32]:

```
model_bias1.compile(optimizer='adam',
                    loss='sparse_categorical_crossentropy',
                    metrics=['sparse_categorical_accuracy'])
```

In [33]:

```
history_bias1 = model_bias1.fit(x_train, y_train,
                                batch_size=128,
                                epochs=100,
                                validation_data=(x_test, y_test),
                                verbose=2)
```

```
Epoch 1/100
469/469 - 1s - loss: 3.1852 - sparse_categorical_accuracy: 0.1880 - val_loss: 1.9531 - val_sparse_categorical_accuracy: 0.2437 - 1s/epoch - 2ms/step
Epoch 2/100
469/469 - 1s - loss: 1.6994 - sparse_categorical_accuracy: 0.3581 - val_loss: 1.4599 - val_sparse_categorical_accuracy: 0.4383 - 808ms/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 1.3357 - sparse_categorical_accuracy: 0.4870 - val_loss: 1.2558 - val_sparse_categorical_accuracy: 0.5083 - 771ms/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 1.1341 - sparse_categorical_accuracy: 0.5530 - val_loss: 1.0741 - val_sparse_categorical_accuracy: 0.5804 - 791ms/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 0.9712 - sparse_categorical_accuracy: 0.6324 - val_loss: 0.9289 - val_sparse_categorical_accuracy: 0.6553 - 759ms/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 0.8669 - sparse_categorical_accuracy: 0.6864 - val_loss: 0.8796 - val_sparse_categorical_accuracy: 0.6896 - 702ms/epoch - 1ms/step
Epoch 7/100
469/469 - 1s - loss: 0.8169 - sparse_categorical_accuracy: 0.7057 - val_loss: 0.8697 - val_sparse_categorical_accuracy: 0.6827 - 681ms/epoch - 1ms/step
Epoch 8/100
469/469 - 1s - loss: 0.7860 - sparse_categorical_accuracy: 0.7191 - val_loss: 0.8116 - val_sparse_categorical_accuracy: 0.7139 - 709ms/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 0.7695 - sparse_categorical_accuracy: 0.7266 - val_loss: 0.7961 - val_sparse_categorical_accuracy: 0.7182 - 806ms/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 0.7530 - sparse_categorical_accuracy: 0.7291 - val_loss: 0.8058 - val_sparse_categorical_accuracy: 0.7188 - 708ms/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 0.7467 - sparse_categorical_accuracy: 0.7311 - val_loss: 0.7823 - val_sparse_categorical_accuracy: 0.7322 - 790ms/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.7250 - sparse_categorical_accuracy: 0.7402 - val_loss: 0.8169 - val_sparse_categorical_accuracy: 0.7402 - 790ms/epoch - 2ms/step
```

tegorical accuracy: 0.7038 - 709ms/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.7096 - sparse_categorical_accuracy: 0.7441 - val_loss: 0.8539 - val_sparse_categorical_accuracy: 0.7132 - 740ms/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.7167 - sparse_categorical_accuracy: 0.7437 - val_loss: 0.7588 - val_sparse_categorical_accuracy: 0.7313 - 698ms/epoch - 1ms/step
Epoch 15/100
469/469 - 1s - loss: 0.6990 - sparse_categorical_accuracy: 0.7481 - val_loss: 0.7695 - val_sparse_categorical_accuracy: 0.7340 - 820ms/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.6906 - sparse_categorical_accuracy: 0.7510 - val_loss: 0.7563 - val_sparse_categorical_accuracy: 0.7269 - 732ms/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.6938 - sparse_categorical_accuracy: 0.7469 - val_loss: 0.7579 - val_sparse_categorical_accuracy: 0.7211 - 726ms/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 0.6781 - sparse_categorical_accuracy: 0.7530 - val_loss: 0.7434 - val_sparse_categorical_accuracy: 0.7330 - 720ms/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.6832 - sparse_categorical_accuracy: 0.7508 - val_loss: 0.7434 - val_sparse_categorical_accuracy: 0.7375 - 753ms/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.6710 - sparse_categorical_accuracy: 0.7555 - val_loss: 0.7531 - val_sparse_categorical_accuracy: 0.7395 - 758ms/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.6722 - sparse_categorical_accuracy: 0.7543 - val_loss: 0.7219 - val_sparse_categorical_accuracy: 0.7319 - 775ms/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.6634 - sparse_categorical_accuracy: 0.7573 - val_loss: 0.7194 - val_sparse_categorical_accuracy: 0.7485 - 791ms/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.6542 - sparse_categorical_accuracy: 0.7615 - val_loss: 0.7349 - val_sparse_categorical_accuracy: 0.7500 - 802ms/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.6471 - sparse_categorical_accuracy: 0.7628 - val_loss: 0.6975 - val_sparse_categorical_accuracy: 0.7552 - 827ms/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.6454 - sparse_categorical_accuracy: 0.7634 - val_loss: 0.6978 - val_sparse_categorical_accuracy: 0.7539 - 774ms/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.6503 - sparse_categorical_accuracy: 0.7607 - val_loss: 0.7485 - val_sparse_categorical_accuracy: 0.7341 - 798ms/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.6458 - sparse_categorical_accuracy: 0.7625 - val_loss: 0.7436 - val_sparse_categorical_accuracy: 0.7440 - 753ms/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.6395 - sparse_categorical_accuracy: 0.7645 - val_loss: 0.7189 - val_sparse_categorical_accuracy: 0.7388 - 807ms/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.6347 - sparse_categorical_accuracy: 0.7660 - val_loss: 0.7296 - val_sparse_categorical_accuracy: 0.7286 - 775ms/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.6382 - sparse_categorical_accuracy: 0.7639 - val_loss: 0.6846 - val_sparse_categorical_accuracy: 0.7633 - 733ms/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.6283 - sparse_categorical_accuracy: 0.7707 - val_loss: 0.6928 - val_sparse_categorical_accuracy: 0.7594 - 798ms/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.6292 - sparse_categorical_accuracy: 0.7680 - val_loss: 0.7008 - val_sparse_categorical_accuracy: 0.7604 - 742ms/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.6413 - sparse_categorical_accuracy: 0.7649 - val_loss: 0.7749 - val_sparse_categorical_accuracy: 0.7185 - 781ms/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.6310 - sparse_categorical_accuracy: 0.7680 - val_loss: 0.7040 - val_sparse_categorical_accuracy: 0.7430 - 767ms/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.6323 - sparse_categorical_accuracy: 0.7647 - val_loss: 0.6631 - val_sparse_categorical_accuracy: 0.7588 - 812ms/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.6262 - sparse_categorical_accuracy: 0.7667 - val_loss: 0.7823 - val_sparse_categorical_accuracy: 0.7113 - 829ms/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.6267 - sparse_categorical_accuracy: 0.7676 - val_loss: 0.7096 - val_sparse_categorical_accuracy: 0.7535 - 782ms/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.6262 - sparse_categorical_accuracy: 0.7692 - val_loss: 0.7149 - val_sparse_categorical_accuracy: 0.7495 - 754ms/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.6225 - sparse_categorical_accuracy: 0.7676 - val_loss: 0.6897 - val_sparse_categorical_accuracy: 0.7598 - 693ms/epoch - 1ms/step
Epoch 40/100

469/469 - 1s - loss: 0.6230 - sparse_categorical_accuracy: 0.7702 - val_loss: 0.7223 - val_sparse_categorical_accuracy: 0.7475 - 754ms/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.6323 - sparse_categorical_accuracy: 0.7654 - val_loss: 0.6992 - val_sparse_categorical_accuracy: 0.7530 - 814ms/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.6224 - sparse_categorical_accuracy: 0.7653 - val_loss: 0.6675 - val_sparse_categorical_accuracy: 0.7579 - 808ms/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.6265 - sparse_categorical_accuracy: 0.7688 - val_loss: 0.6822 - val_sparse_categorical_accuracy: 0.7524 - 799ms/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.6251 - sparse_categorical_accuracy: 0.7675 - val_loss: 0.6763 - val_sparse_categorical_accuracy: 0.7542 - 719ms/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.6200 - sparse_categorical_accuracy: 0.7707 - val_loss: 0.6694 - val_sparse_categorical_accuracy: 0.7627 - 783ms/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.6163 - sparse_categorical_accuracy: 0.7703 - val_loss: 0.6987 - val_sparse_categorical_accuracy: 0.7425 - 796ms/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.6186 - sparse_categorical_accuracy: 0.7699 - val_loss: 0.6894 - val_sparse_categorical_accuracy: 0.7437 - 822ms/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.6215 - sparse_categorical_accuracy: 0.7714 - val_loss: 0.7270 - val_sparse_categorical_accuracy: 0.7150 - 752ms/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.6247 - sparse_categorical_accuracy: 0.7711 - val_loss: 0.6824 - val_sparse_categorical_accuracy: 0.7643 - 750ms/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.6153 - sparse_categorical_accuracy: 0.7748 - val_loss: 0.6548 - val_sparse_categorical_accuracy: 0.7651 - 814ms/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.6141 - sparse_categorical_accuracy: 0.7748 - val_loss: 0.6912 - val_sparse_categorical_accuracy: 0.7587 - 787ms/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.6146 - sparse_categorical_accuracy: 0.7746 - val_loss: 0.7133 - val_sparse_categorical_accuracy: 0.7565 - 811ms/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.6184 - sparse_categorical_accuracy: 0.7749 - val_loss: 0.7336 - val_sparse_categorical_accuracy: 0.7398 - 789ms/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.6102 - sparse_categorical_accuracy: 0.7734 - val_loss: 0.6870 - val_sparse_categorical_accuracy: 0.7671 - 836ms/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.6218 - sparse_categorical_accuracy: 0.7703 - val_loss: 0.7047 - val_sparse_categorical_accuracy: 0.7400 - 845ms/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.6207 - sparse_categorical_accuracy: 0.7691 - val_loss: 0.6838 - val_sparse_categorical_accuracy: 0.7569 - 754ms/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 0.6093 - sparse_categorical_accuracy: 0.7724 - val_loss: 0.6662 - val_sparse_categorical_accuracy: 0.7669 - 797ms/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.6047 - sparse_categorical_accuracy: 0.7781 - val_loss: 0.7107 - val_sparse_categorical_accuracy: 0.7558 - 779ms/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.6094 - sparse_categorical_accuracy: 0.7726 - val_loss: 0.6969 - val_sparse_categorical_accuracy: 0.7415 - 746ms/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.6111 - sparse_categorical_accuracy: 0.7729 - val_loss: 0.6583 - val_sparse_categorical_accuracy: 0.7654 - 748ms/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.6088 - sparse_categorical_accuracy: 0.7797 - val_loss: 0.6702 - val_sparse_categorical_accuracy: 0.7641 - 763ms/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.6122 - sparse_categorical_accuracy: 0.7704 - val_loss: 0.6719 - val_sparse_categorical_accuracy: 0.7339 - 703ms/epoch - 1ms/step
Epoch 63/100
469/469 - 1s - loss: 0.5989 - sparse_categorical_accuracy: 0.7782 - val_loss: 0.6558 - val_sparse_categorical_accuracy: 0.7670 - 783ms/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.6108 - sparse_categorical_accuracy: 0.7733 - val_loss: 0.6806 - val_sparse_categorical_accuracy: 0.7621 - 706ms/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.6143 - sparse_categorical_accuracy: 0.7730 - val_loss: 0.6871 - val_sparse_categorical_accuracy: 0.7425 - 727ms/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.6031 - sparse_categorical_accuracy: 0.7763 - val_loss: 0.6662 - val_sparse_categorical_accuracy: 0.7573 - 742ms/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.6060 - sparse_categorical_accuracy: 0.7728 - val_loss: 0.6740 - val_sparse_categorical_accuracy: 0.7659 - 804ms/epoch - 2ms/step

Epoch 68/100
469/469 - 1s - loss: 0.6027 - sparse_categorical_accuracy: 0.7780 - val_loss: 0.6667 - val_sparse_categorical_accuracy: 0.7641 - 806ms/epoch - 2ms/step

Epoch 69/100
469/469 - 1s - loss: 0.6026 - sparse_categorical_accuracy: 0.7766 - val_loss: 0.6596 - val_sparse_categorical_accuracy: 0.7651 - 772ms/epoch - 2ms/step

Epoch 70/100
469/469 - 1s - loss: 0.6093 - sparse_categorical_accuracy: 0.7739 - val_loss: 0.6533 - val_sparse_categorical_accuracy: 0.7581 - 794ms/epoch - 2ms/step

Epoch 71/100
469/469 - 1s - loss: 0.6121 - sparse_categorical_accuracy: 0.7724 - val_loss: 0.6456 - val_sparse_categorical_accuracy: 0.7611 - 737ms/epoch - 2ms/step

Epoch 72/100
469/469 - 1s - loss: 0.5903 - sparse_categorical_accuracy: 0.7790 - val_loss: 0.6458 - val_sparse_categorical_accuracy: 0.7630 - 788ms/epoch - 2ms/step

Epoch 73/100
469/469 - 1s - loss: 0.6089 - sparse_categorical_accuracy: 0.7706 - val_loss: 0.6503 - val_sparse_categorical_accuracy: 0.7699 - 835ms/epoch - 2ms/step

Epoch 74/100
469/469 - 1s - loss: 0.6030 - sparse_categorical_accuracy: 0.7753 - val_loss: 0.6578 - val_sparse_categorical_accuracy: 0.7544 - 783ms/epoch - 2ms/step

Epoch 75/100
469/469 - 1s - loss: 0.5921 - sparse_categorical_accuracy: 0.7789 - val_loss: 0.6539 - val_sparse_categorical_accuracy: 0.7651 - 779ms/epoch - 2ms/step

Epoch 76/100
469/469 - 1s - loss: 0.6033 - sparse_categorical_accuracy: 0.7731 - val_loss: 0.6674 - val_sparse_categorical_accuracy: 0.7661 - 775ms/epoch - 2ms/step

Epoch 77/100
469/469 - 1s - loss: 0.6006 - sparse_categorical_accuracy: 0.7812 - val_loss: 0.6788 - val_sparse_categorical_accuracy: 0.7557 - 853ms/epoch - 2ms/step

Epoch 78/100
469/469 - 1s - loss: 0.5964 - sparse_categorical_accuracy: 0.7759 - val_loss: 0.6649 - val_sparse_categorical_accuracy: 0.7491 - 774ms/epoch - 2ms/step

Epoch 79/100
469/469 - 1s - loss: 0.5896 - sparse_categorical_accuracy: 0.7800 - val_loss: 0.6418 - val_sparse_categorical_accuracy: 0.7681 - 717ms/epoch - 2ms/step

Epoch 80/100
469/469 - 1s - loss: 0.6131 - sparse_categorical_accuracy: 0.7753 - val_loss: 0.6679 - val_sparse_categorical_accuracy: 0.7518 - 718ms/epoch - 2ms/step

Epoch 81/100
469/469 - 1s - loss: 0.5913 - sparse_categorical_accuracy: 0.7768 - val_loss: 0.6524 - val_sparse_categorical_accuracy: 0.7706 - 755ms/epoch - 2ms/step

Epoch 82/100
469/469 - 1s - loss: 0.6011 - sparse_categorical_accuracy: 0.7767 - val_loss: 0.6847 - val_sparse_categorical_accuracy: 0.7279 - 758ms/epoch - 2ms/step

Epoch 83/100
469/469 - 1s - loss: 0.5774 - sparse_categorical_accuracy: 0.7863 - val_loss: 0.6568 - val_sparse_categorical_accuracy: 0.7726 - 731ms/epoch - 2ms/step

Epoch 84/100
469/469 - 1s - loss: 0.5854 - sparse_categorical_accuracy: 0.7811 - val_loss: 0.6583 - val_sparse_categorical_accuracy: 0.7693 - 779ms/epoch - 2ms/step

Epoch 85/100
469/469 - 1s - loss: 0.5848 - sparse_categorical_accuracy: 0.7818 - val_loss: 0.6646 - val_sparse_categorical_accuracy: 0.7650 - 787ms/epoch - 2ms/step

Epoch 86/100
469/469 - 1s - loss: 0.5901 - sparse_categorical_accuracy: 0.7804 - val_loss: 0.6404 - val_sparse_categorical_accuracy: 0.7678 - 801ms/epoch - 2ms/step

Epoch 87/100
469/469 - 1s - loss: 0.5907 - sparse_categorical_accuracy: 0.7812 - val_loss: 0.7260 - val_sparse_categorical_accuracy: 0.7299 - 731ms/epoch - 2ms/step

Epoch 88/100
469/469 - 1s - loss: 0.5940 - sparse_categorical_accuracy: 0.7747 - val_loss: 0.6579 - val_sparse_categorical_accuracy: 0.7557 - 720ms/epoch - 2ms/step

Epoch 89/100
469/469 - 1s - loss: 0.5819 - sparse_categorical_accuracy: 0.7816 - val_loss: 0.6369 - val_sparse_categorical_accuracy: 0.7638 - 824ms/epoch - 2ms/step

Epoch 90/100
469/469 - 1s - loss: 0.5943 - sparse_categorical_accuracy: 0.7754 - val_loss: 0.6341 - val_sparse_categorical_accuracy: 0.7751 - 789ms/epoch - 2ms/step

Epoch 91/100
469/469 - 1s - loss: 0.5816 - sparse_categorical_accuracy: 0.7815 - val_loss: 0.6377 - val_sparse_categorical_accuracy: 0.7637 - 795ms/epoch - 2ms/step

Epoch 92/100
469/469 - 1s - loss: 0.5905 - sparse_categorical_accuracy: 0.7796 - val_loss: 0.6547 - val_sparse_categorical_accuracy: 0.7658 - 704ms/epoch - 2ms/step

Epoch 93/100
469/469 - 1s - loss: 0.5798 - sparse_categorical_accuracy: 0.7811 - val_loss: 0.6274 - val_sparse_categorical_accuracy: 0.7674 - 820ms/epoch - 2ms/step

Epoch 94/100
469/469 - 1s - loss: 0.5741 - sparse_categorical_accuracy: 0.7842 - val_loss: 0.6580 - val_sparse_categorical_accuracy: 0.7681 - 728ms/epoch - 2ms/step

Epoch 95/100
469/469 - 1s - loss: 0.5893 - sparse_categorical_accuracy: 0.7777 - val_loss: 0.6977 - val_sparse_categorical_accuracy: 0.7681 - 728ms/epoch - 2ms/step

```
tegorical_accuracy: 0.7149 - 821ms/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.5759 - sparse_categorical_accuracy: 0.7812 - val_loss: 0.6679 - val_sparse_ca
tegorical_accuracy: 0.7611 - 788ms/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.5716 - sparse_categorical_accuracy: 0.7890 - val_loss: 0.6473 - val_sparse_ca
tegorical_accuracy: 0.7726 - 833ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.5760 - sparse_categorical_accuracy: 0.7863 - val_loss: 0.6454 - val_sparse_ca
tegorical_accuracy: 0.7740 - 805ms/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.5845 - sparse_categorical_accuracy: 0.7811 - val_loss: 0.6730 - val_sparse_ca
tegorical_accuracy: 0.7665 - 729ms/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.5834 - sparse_categorical_accuracy: 0.7813 - val_loss: 0.6432 - val_sparse_ca
tegorical_accuracy: 0.7701 - 723ms/epoch - 2ms/step
```

In [34]:

```
model_bias1.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.6432 - sparse_categorical_accuracy: 0.7701 - 235ms/epoch - 752us/step
```

Out[34]:

```
[0.6431722044944763, 0.7700999975204468]
```

Model with Ones bias initializer

In [35]:

```
model_bias2 = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(16, activation='relu', bias_initializer='Ones'),
    tf.keras.layers.Dense(16, activation='relu', bias_initializer='Ones'),
    tf.keras.layers.Dense(10, activation='softmax', bias_initializer='Ones')
])
```

In [36]:

```
model_bias2.compile(optimizer='adam',
                    loss='sparse_categorical_crossentropy',
                    metrics=['sparse_categorical_accuracy'])
```

In [37]:

```
history_bias2 = model_bias2.fit(x_train, y_train,
                                batch_size=128,
                                epochs=100,
                                validation_data=(x_test, y_test),
                                verbose=2)
```

```
Epoch 1/100
469/469 - 1s - loss: 3.4335 - sparse_categorical_accuracy: 0.2413 - val_loss: 1.7527 - val_sparse_ca
tegorical_accuracy: 0.3882 - 1s/epoch - 3ms/step
Epoch 2/100
469/469 - 1s - loss: 1.5316 - sparse_categorical_accuracy: 0.4762 - val_loss: 1.4887 - val_sparse_ca
tegorical_accuracy: 0.4959 - 839ms/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 1.3172 - sparse_categorical_accuracy: 0.5413 - val_loss: 1.2508 - val_sparse_ca
tegorical_accuracy: 0.5636 - 760ms/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 1.1362 - sparse_categorical_accuracy: 0.6036 - val_loss: 1.0150 - val_sparse_ca
tegorical_accuracy: 0.6878 - 828ms/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 0.8998 - sparse_categorical_accuracy: 0.7082 - val_loss: 0.8755 - val_sparse_ca
tegorical_accuracy: 0.7194 - 812ms/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 0.8141 - sparse_categorical_accuracy: 0.7270 - val_loss: 0.8025 - val_sparse_ca
tegorical_accuracy: 0.7366 - 785ms/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.7692 - sparse_categorical_accuracy: 0.7393 - val_loss: 0.7780 - val_sparse_ca
tegorical_accuracy: 0.7410 - 792ms/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 0.7325 - sparse_categorical_accuracy: 0.7444 - val_loss: 0.7696 - val_sparse_ca
tegorical_accuracy: 0.7422 - 779ms/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 0.7164 - sparse_categorical_accuracy: 0.7491 - val_loss: 0.7428 - val_sparse_ca
tegorical_accuracy: 0.7538 - 731ms/epoch - 2ms/step
Epoch 10/100
```

469/469 - 1s - loss: 0.6946 - sparse_categorical_accuracy: 0.7550 - val_loss: 0.7790 - val_sparse_categorical_accuracy: 0.7371 - 798ms/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 0.6834 - sparse_categorical_accuracy: 0.7578 - val_loss: 0.7125 - val_sparse_categorical_accuracy: 0.7521 - 826ms/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.6693 - sparse_categorical_accuracy: 0.7615 - val_loss: 0.7316 - val_sparse_categorical_accuracy: 0.7490 - 762ms/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.6700 - sparse_categorical_accuracy: 0.7601 - val_loss: 0.7199 - val_sparse_categorical_accuracy: 0.7534 - 759ms/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.6585 - sparse_categorical_accuracy: 0.7628 - val_loss: 0.6997 - val_sparse_categorical_accuracy: 0.7539 - 736ms/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 0.6492 - sparse_categorical_accuracy: 0.7653 - val_loss: 0.7297 - val_sparse_categorical_accuracy: 0.7527 - 755ms/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.6506 - sparse_categorical_accuracy: 0.7655 - val_loss: 0.7319 - val_sparse_categorical_accuracy: 0.7455 - 812ms/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.6417 - sparse_categorical_accuracy: 0.7666 - val_loss: 0.7004 - val_sparse_categorical_accuracy: 0.7536 - 813ms/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 0.6436 - sparse_categorical_accuracy: 0.7666 - val_loss: 0.7002 - val_sparse_categorical_accuracy: 0.7590 - 769ms/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.6329 - sparse_categorical_accuracy: 0.7698 - val_loss: 0.6925 - val_sparse_categorical_accuracy: 0.7600 - 763ms/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.6315 - sparse_categorical_accuracy: 0.7709 - val_loss: 0.6737 - val_sparse_categorical_accuracy: 0.7623 - 753ms/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.6286 - sparse_categorical_accuracy: 0.7708 - val_loss: 0.6812 - val_sparse_categorical_accuracy: 0.7601 - 733ms/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.6260 - sparse_categorical_accuracy: 0.7712 - val_loss: 0.7359 - val_sparse_categorical_accuracy: 0.7406 - 829ms/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.6215 - sparse_categorical_accuracy: 0.7728 - val_loss: 0.6720 - val_sparse_categorical_accuracy: 0.7657 - 740ms/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.6232 - sparse_categorical_accuracy: 0.7706 - val_loss: 0.6852 - val_sparse_categorical_accuracy: 0.7563 - 784ms/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.6179 - sparse_categorical_accuracy: 0.7742 - val_loss: 0.6643 - val_sparse_categorical_accuracy: 0.7670 - 740ms/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.6145 - sparse_categorical_accuracy: 0.7743 - val_loss: 0.6745 - val_sparse_categorical_accuracy: 0.7598 - 827ms/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.6114 - sparse_categorical_accuracy: 0.7765 - val_loss: 0.7063 - val_sparse_categorical_accuracy: 0.7545 - 797ms/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.6049 - sparse_categorical_accuracy: 0.7790 - val_loss: 0.6690 - val_sparse_categorical_accuracy: 0.7649 - 743ms/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.6033 - sparse_categorical_accuracy: 0.7781 - val_loss: 0.6601 - val_sparse_categorical_accuracy: 0.7671 - 802ms/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.6021 - sparse_categorical_accuracy: 0.7770 - val_loss: 0.7091 - val_sparse_categorical_accuracy: 0.7482 - 785ms/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.6089 - sparse_categorical_accuracy: 0.7752 - val_loss: 0.7153 - val_sparse_categorical_accuracy: 0.7470 - 741ms/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.6097 - sparse_categorical_accuracy: 0.7772 - val_loss: 0.6604 - val_sparse_categorical_accuracy: 0.7658 - 815ms/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.5971 - sparse_categorical_accuracy: 0.7813 - val_loss: 0.6549 - val_sparse_categorical_accuracy: 0.7645 - 779ms/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.5946 - sparse_categorical_accuracy: 0.7797 - val_loss: 0.6444 - val_sparse_categorical_accuracy: 0.7693 - 820ms/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.5873 - sparse_categorical_accuracy: 0.7786 - val_loss: 0.7113 - val_sparse_categorical_accuracy: 0.7380 - 853ms/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.5852 - sparse_categorical_accuracy: 0.7801 - val_loss: 0.7121 - val_sparse_categorical_accuracy: 0.7551 - 801ms/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.5777 - sparse_categorical_accuracy: 0.7824 - val_loss: 0.6408 - val_sparse_categorical_accuracy: 0.7737 - 710ms/epoch - 2ms/step

Epoch 38/100
469/469 - 1s - loss: 0.5753 - sparse_categorical_accuracy: 0.7834 - val_loss: 0.6409 - val_sparse_categorical_accuracy: 0.7687 - 816ms/epoch - 2ms/step

Epoch 39/100
469/469 - 1s - loss: 0.5758 - sparse_categorical_accuracy: 0.7828 - val_loss: 0.6333 - val_sparse_categorical_accuracy: 0.7681 - 809ms/epoch - 2ms/step

Epoch 40/100
469/469 - 1s - loss: 0.5711 - sparse_categorical_accuracy: 0.7847 - val_loss: 0.6488 - val_sparse_categorical_accuracy: 0.7659 - 736ms/epoch - 2ms/step

Epoch 41/100
469/469 - 1s - loss: 0.5663 - sparse_categorical_accuracy: 0.7867 - val_loss: 0.6282 - val_sparse_categorical_accuracy: 0.7756 - 717ms/epoch - 2ms/step

Epoch 42/100
469/469 - 1s - loss: 0.5633 - sparse_categorical_accuracy: 0.7882 - val_loss: 0.6393 - val_sparse_categorical_accuracy: 0.7716 - 800ms/epoch - 2ms/step

Epoch 43/100
469/469 - 1s - loss: 0.5613 - sparse_categorical_accuracy: 0.7877 - val_loss: 0.6228 - val_sparse_categorical_accuracy: 0.7766 - 746ms/epoch - 2ms/step

Epoch 44/100
469/469 - 1s - loss: 0.5593 - sparse_categorical_accuracy: 0.7906 - val_loss: 0.6253 - val_sparse_categorical_accuracy: 0.7753 - 758ms/epoch - 2ms/step

Epoch 45/100
469/469 - 1s - loss: 0.5613 - sparse_categorical_accuracy: 0.7896 - val_loss: 0.6212 - val_sparse_categorical_accuracy: 0.7765 - 814ms/epoch - 2ms/step

Epoch 46/100
469/469 - 1s - loss: 0.5375 - sparse_categorical_accuracy: 0.7977 - val_loss: 0.6111 - val_sparse_categorical_accuracy: 0.7791 - 725ms/epoch - 2ms/step

Epoch 47/100
469/469 - 1s - loss: 0.5420 - sparse_categorical_accuracy: 0.7942 - val_loss: 0.6144 - val_sparse_categorical_accuracy: 0.7767 - 835ms/epoch - 2ms/step

Epoch 48/100
469/469 - 1s - loss: 0.5395 - sparse_categorical_accuracy: 0.7958 - val_loss: 0.6248 - val_sparse_categorical_accuracy: 0.7795 - 823ms/epoch - 2ms/step

Epoch 49/100
469/469 - 1s - loss: 0.5367 - sparse_categorical_accuracy: 0.7981 - val_loss: 0.6159 - val_sparse_categorical_accuracy: 0.7781 - 761ms/epoch - 2ms/step

Epoch 50/100
469/469 - 1s - loss: 0.5288 - sparse_categorical_accuracy: 0.8005 - val_loss: 0.6189 - val_sparse_categorical_accuracy: 0.7807 - 705ms/epoch - 2ms/step

Epoch 51/100
469/469 - 1s - loss: 0.5283 - sparse_categorical_accuracy: 0.7994 - val_loss: 0.6309 - val_sparse_categorical_accuracy: 0.7688 - 715ms/epoch - 2ms/step

Epoch 52/100
469/469 - 1s - loss: 0.5262 - sparse_categorical_accuracy: 0.7995 - val_loss: 0.5984 - val_sparse_categorical_accuracy: 0.7809 - 779ms/epoch - 2ms/step

Epoch 53/100
469/469 - 1s - loss: 0.5259 - sparse_categorical_accuracy: 0.8008 - val_loss: 0.5985 - val_sparse_categorical_accuracy: 0.7908 - 793ms/epoch - 2ms/step

Epoch 54/100
469/469 - 1s - loss: 0.5180 - sparse_categorical_accuracy: 0.8051 - val_loss: 0.5828 - val_sparse_categorical_accuracy: 0.7888 - 728ms/epoch - 2ms/step

Epoch 55/100
469/469 - 1s - loss: 0.5171 - sparse_categorical_accuracy: 0.8033 - val_loss: 0.5984 - val_sparse_categorical_accuracy: 0.7846 - 786ms/epoch - 2ms/step

Epoch 56/100
469/469 - 1s - loss: 0.5171 - sparse_categorical_accuracy: 0.8034 - val_loss: 0.6036 - val_sparse_categorical_accuracy: 0.7902 - 753ms/epoch - 2ms/step

Epoch 57/100
469/469 - 1s - loss: 0.5133 - sparse_categorical_accuracy: 0.8044 - val_loss: 0.6113 - val_sparse_categorical_accuracy: 0.7824 - 782ms/epoch - 2ms/step

Epoch 58/100
469/469 - 1s - loss: 0.5115 - sparse_categorical_accuracy: 0.8047 - val_loss: 0.5941 - val_sparse_categorical_accuracy: 0.7902 - 732ms/epoch - 2ms/step

Epoch 59/100
469/469 - 1s - loss: 0.5053 - sparse_categorical_accuracy: 0.8051 - val_loss: 0.5757 - val_sparse_categorical_accuracy: 0.7938 - 777ms/epoch - 2ms/step

Epoch 60/100
469/469 - 1s - loss: 0.5053 - sparse_categorical_accuracy: 0.8072 - val_loss: 0.6414 - val_sparse_categorical_accuracy: 0.7676 - 747ms/epoch - 2ms/step

Epoch 61/100
469/469 - 1s - loss: 0.5100 - sparse_categorical_accuracy: 0.8053 - val_loss: 0.5995 - val_sparse_categorical_accuracy: 0.7952 - 769ms/epoch - 2ms/step

Epoch 62/100
469/469 - 1s - loss: 0.5026 - sparse_categorical_accuracy: 0.8084 - val_loss: 0.5706 - val_sparse_categorical_accuracy: 0.7958 - 758ms/epoch - 2ms/step

Epoch 63/100
469/469 - 1s - loss: 0.4991 - sparse_categorical_accuracy: 0.8100 - val_loss: 0.5873 - val_sparse_categorical_accuracy: 0.7936 - 784ms/epoch - 2ms/step

Epoch 64/100
469/469 - 1s - loss: 0.5004 - sparse_categorical_accuracy: 0.8085 - val_loss: 0.5985 - val_sparse_categorical_accuracy: 0.7852 - 847ms/epoch - 2ms/step

Epoch 65/100
469/469 - 1s - loss: 0.5030 - sparse_categorical_accuracy: 0.8097 - val_loss: 0.5889 - val_sparse_categorical_accuracy: 0.7852 - 847ms/epoch - 2ms/step

tegorical accuracy: 0.7928 - 782ms/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.4973 - sparse_categorical_accuracy: 0.8102 - val_loss: 0.5877 - val_sparse_ca
tegorical accuracy: 0.7942 - 797ms/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.4986 - sparse_categorical_accuracy: 0.8102 - val_loss: 0.5746 - val_sparse_ca
tegorical accuracy: 0.8016 - 816ms/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 0.4900 - sparse_categorical_accuracy: 0.8141 - val_loss: 0.5798 - val_sparse_ca
tegorical accuracy: 0.8012 - 787ms/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 0.4954 - sparse_categorical_accuracy: 0.8104 - val_loss: 0.5717 - val_sparse_ca
tegorical accuracy: 0.7999 - 800ms/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.4975 - sparse_categorical_accuracy: 0.8108 - val_loss: 0.5714 - val_sparse_ca
tegorical accuracy: 0.7965 - 761ms/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.4899 - sparse_categorical_accuracy: 0.8129 - val_loss: 0.5918 - val_sparse_ca
tegorical accuracy: 0.7944 - 819ms/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.4849 - sparse_categorical_accuracy: 0.8159 - val_loss: 0.5749 - val_sparse_ca
tegorical accuracy: 0.8011 - 771ms/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.4867 - sparse_categorical_accuracy: 0.8145 - val_loss: 0.5856 - val_sparse_ca
tegorical accuracy: 0.7977 - 735ms/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.4856 - sparse_categorical_accuracy: 0.8146 - val_loss: 0.5758 - val_sparse_ca
tegorical accuracy: 0.8001 - 839ms/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.4793 - sparse_categorical_accuracy: 0.8169 - val_loss: 0.5993 - val_sparse_ca
tegorical accuracy: 0.7867 - 731ms/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.4828 - sparse_categorical_accuracy: 0.8170 - val_loss: 0.5935 - val_sparse_ca
tegorical accuracy: 0.8009 - 789ms/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.4854 - sparse_categorical_accuracy: 0.8163 - val_loss: 0.5770 - val_sparse_ca
tegorical accuracy: 0.7965 - 759ms/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.4791 - sparse_categorical_accuracy: 0.8170 - val_loss: 0.5655 - val_sparse_ca
tegorical accuracy: 0.8025 - 767ms/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.4793 - sparse_categorical_accuracy: 0.8173 - val_loss: 0.5685 - val_sparse_ca
tegorical accuracy: 0.8065 - 764ms/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.4770 - sparse_categorical_accuracy: 0.8212 - val_loss: 0.5744 - val_sparse_ca
tegorical accuracy: 0.8025 - 810ms/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.4734 - sparse_categorical_accuracy: 0.8217 - val_loss: 0.5755 - val_sparse_ca
tegorical accuracy: 0.8009 - 784ms/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.4680 - sparse_categorical_accuracy: 0.8249 - val_loss: 0.5669 - val_sparse_ca
tegorical accuracy: 0.8085 - 859ms/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.4622 - sparse_categorical_accuracy: 0.8288 - val_loss: 0.5619 - val_sparse_ca
tegorical accuracy: 0.8073 - 748ms/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.4597 - sparse_categorical_accuracy: 0.8282 - val_loss: 0.5644 - val_sparse_ca
tegorical accuracy: 0.8048 - 747ms/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.4535 - sparse_categorical_accuracy: 0.8317 - val_loss: 0.5838 - val_sparse_ca
tegorical accuracy: 0.7979 - 838ms/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.4549 - sparse_categorical_accuracy: 0.8310 - val_loss: 0.5551 - val_sparse_ca
tegorical accuracy: 0.8142 - 787ms/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.4474 - sparse_categorical_accuracy: 0.8333 - val_loss: 0.5547 - val_sparse_ca
tegorical accuracy: 0.8125 - 754ms/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.4462 - sparse_categorical_accuracy: 0.8347 - val_loss: 0.5528 - val_sparse_ca
tegorical accuracy: 0.8049 - 733ms/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.4450 - sparse_categorical_accuracy: 0.8367 - val_loss: 0.5380 - val_sparse_ca
tegorical accuracy: 0.8178 - 754ms/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.4424 - sparse_categorical_accuracy: 0.8357 - val_loss: 0.5295 - val_sparse_ca
tegorical accuracy: 0.8198 - 710ms/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.4387 - sparse_categorical_accuracy: 0.8356 - val_loss: 0.5336 - val_sparse_ca
tegorical accuracy: 0.8166 - 780ms/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.4394 - sparse_categorical_accuracy: 0.8369 - val_loss: 0.5528 - val_sparse_ca
tegorical accuracy: 0.8158 - 825ms/epoch - 2ms/step
Epoch 93/100

```
469/469 - 1s - loss: 0.4443 - sparse_categorical_accuracy: 0.8365 - val_loss: 0.5343 - val_sparse_categorical_accuracy: 0.8129 - 789ms/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.4422 - sparse_categorical_accuracy: 0.8374 - val_loss: 0.5701 - val_sparse_categorical_accuracy: 0.8002 - 821ms/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.4346 - sparse_categorical_accuracy: 0.8382 - val_loss: 0.5302 - val_sparse_categorical_accuracy: 0.8219 - 822ms/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.4364 - sparse_categorical_accuracy: 0.8395 - val_loss: 0.5278 - val_sparse_categorical_accuracy: 0.8193 - 674ms/epoch - 1ms/step
Epoch 97/100
469/469 - 1s - loss: 0.4345 - sparse_categorical_accuracy: 0.8391 - val_loss: 0.5322 - val_sparse_categorical_accuracy: 0.8177 - 765ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.4356 - sparse_categorical_accuracy: 0.8397 - val_loss: 0.5603 - val_sparse_categorical_accuracy: 0.8076 - 824ms/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.4358 - sparse_categorical_accuracy: 0.8390 - val_loss: 0.5267 - val_sparse_categorical_accuracy: 0.8171 - 740ms/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.4314 - sparse_categorical_accuracy: 0.8409 - val_loss: 0.5359 - val_sparse_categorical_accuracy: 0.8148 - 807ms/epoch - 2ms/step
```

In [38]:

```
model_bias2.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.5359 - sparse_categorical_accuracy: 0.8148 - 236ms/epoch - 755us/step
```

Out[38]:

```
[0.5358713269233704, 0.8148000240325928]
```

Model with random_uniform bias initializer

In [39]:

```
model_bias3 = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(16, activation='relu', bias_initializer='random_uniform'),
    tf.keras.layers.Dense(16, activation='relu', bias_initializer='random_uniform'),
    tf.keras.layers.Dense(10, activation='softmax', bias_initializer='random_uniform')
])
```

In [40]:

```
model_bias3.compile(optimizer='adam',
                    loss='sparse_categorical_crossentropy',
                    metrics=['sparse_categorical_accuracy'])
```

In [41]:

```
history_bias3 = model_bias3.fit(x_train, y_train,
                                batch_size=128,
                                epochs=100,
                                validation_data=(x_test, y_test),
                                verbose=2)
```

```
Epoch 1/100
469/469 - 1s - loss: 2.8951 - sparse_categorical_accuracy: 0.1112 - val_loss: 2.2337 - val_sparse_categorical_accuracy: 0.1376 - 1s/epoch - 3ms/step
Epoch 2/100
469/469 - 1s - loss: 2.1250 - sparse_categorical_accuracy: 0.1776 - val_loss: 2.0125 - val_sparse_categorical_accuracy: 0.1936 - 859ms/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 1.9052 - sparse_categorical_accuracy: 0.2553 - val_loss: 1.6792 - val_sparse_categorical_accuracy: 0.3419 - 805ms/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 1.5506 - sparse_categorical_accuracy: 0.3927 - val_loss: 1.4620 - val_sparse_categorical_accuracy: 0.4097 - 857ms/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 1.3642 - sparse_categorical_accuracy: 0.4438 - val_loss: 1.3462 - val_sparse_categorical_accuracy: 0.4535 - 761ms/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 1.3055 - sparse_categorical_accuracy: 0.4511 - val_loss: 1.2982 - val_sparse_categorical_accuracy: 0.4507 - 921ms/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 1.2696 - sparse_categorical_accuracy: 0.4557 - val_loss: 1.2637 - val_sparse_categorical_accuracy: 0.4600 - 733ms/epoch - 2ms/step
```

Epoch 8/100
469/469 - 1s - loss: 1.2452 - sparse_categorical_accuracy: 0.4615 - val_loss: 1.2415 - val_sparse_categorical_accuracy: 0.4606 - 836ms/epoch - 2ms/step

Epoch 9/100
469/469 - 1s - loss: 1.2282 - sparse_categorical_accuracy: 0.4625 - val_loss: 1.2330 - val_sparse_categorical_accuracy: 0.4658 - 818ms/epoch - 2ms/step

Epoch 10/100
469/469 - 1s - loss: 1.2079 - sparse_categorical_accuracy: 0.4747 - val_loss: 1.1702 - val_sparse_categorical_accuracy: 0.5134 - 763ms/epoch - 2ms/step

Epoch 11/100
469/469 - 1s - loss: 1.0788 - sparse_categorical_accuracy: 0.5405 - val_loss: 1.0723 - val_sparse_categorical_accuracy: 0.5525 - 791ms/epoch - 2ms/step

Epoch 12/100
469/469 - 1s - loss: 1.0121 - sparse_categorical_accuracy: 0.5587 - val_loss: 1.0051 - val_sparse_categorical_accuracy: 0.5607 - 787ms/epoch - 2ms/step

Epoch 13/100
469/469 - 1s - loss: 0.9700 - sparse_categorical_accuracy: 0.5778 - val_loss: 0.9924 - val_sparse_categorical_accuracy: 0.5900 - 823ms/epoch - 2ms/step

Epoch 14/100
469/469 - 1s - loss: 0.9407 - sparse_categorical_accuracy: 0.5956 - val_loss: 0.9584 - val_sparse_categorical_accuracy: 0.6302 - 818ms/epoch - 2ms/step

Epoch 15/100
469/469 - 1s - loss: 0.8921 - sparse_categorical_accuracy: 0.6299 - val_loss: 0.9252 - val_sparse_categorical_accuracy: 0.6090 - 780ms/epoch - 2ms/step

Epoch 16/100
469/469 - 1s - loss: 0.8508 - sparse_categorical_accuracy: 0.6412 - val_loss: 0.8683 - val_sparse_categorical_accuracy: 0.6460 - 845ms/epoch - 2ms/step

Epoch 17/100
469/469 - 1s - loss: 0.8360 - sparse_categorical_accuracy: 0.6450 - val_loss: 0.8851 - val_sparse_categorical_accuracy: 0.6397 - 820ms/epoch - 2ms/step

Epoch 18/100
469/469 - 1s - loss: 0.8191 - sparse_categorical_accuracy: 0.6503 - val_loss: 0.8417 - val_sparse_categorical_accuracy: 0.6545 - 768ms/epoch - 2ms/step

Epoch 19/100
469/469 - 1s - loss: 0.8118 - sparse_categorical_accuracy: 0.6509 - val_loss: 0.8150 - val_sparse_categorical_accuracy: 0.6615 - 722ms/epoch - 2ms/step

Epoch 20/100
469/469 - 1s - loss: 0.7992 - sparse_categorical_accuracy: 0.6574 - val_loss: 0.8357 - val_sparse_categorical_accuracy: 0.6656 - 805ms/epoch - 2ms/step

Epoch 21/100
469/469 - 1s - loss: 0.8002 - sparse_categorical_accuracy: 0.6564 - val_loss: 0.8388 - val_sparse_categorical_accuracy: 0.6494 - 802ms/epoch - 2ms/step

Epoch 22/100
469/469 - 1s - loss: 0.7917 - sparse_categorical_accuracy: 0.6580 - val_loss: 0.9622 - val_sparse_categorical_accuracy: 0.5739 - 777ms/epoch - 2ms/step

Epoch 23/100
469/469 - 1s - loss: 0.7894 - sparse_categorical_accuracy: 0.6592 - val_loss: 0.8127 - val_sparse_categorical_accuracy: 0.6643 - 784ms/epoch - 2ms/step

Epoch 24/100
469/469 - 1s - loss: 0.7688 - sparse_categorical_accuracy: 0.6886 - val_loss: 0.7800 - val_sparse_categorical_accuracy: 0.7264 - 715ms/epoch - 2ms/step

Epoch 25/100
469/469 - 1s - loss: 0.7056 - sparse_categorical_accuracy: 0.7476 - val_loss: 0.6871 - val_sparse_categorical_accuracy: 0.7597 - 830ms/epoch - 2ms/step

Epoch 26/100
469/469 - 1s - loss: 0.6567 - sparse_categorical_accuracy: 0.7641 - val_loss: 0.6860 - val_sparse_categorical_accuracy: 0.7588 - 802ms/epoch - 2ms/step

Epoch 27/100
469/469 - 1s - loss: 0.6459 - sparse_categorical_accuracy: 0.7677 - val_loss: 0.6377 - val_sparse_categorical_accuracy: 0.7757 - 836ms/epoch - 2ms/step

Epoch 28/100
469/469 - 1s - loss: 0.6325 - sparse_categorical_accuracy: 0.7730 - val_loss: 0.6415 - val_sparse_categorical_accuracy: 0.7713 - 818ms/epoch - 2ms/step

Epoch 29/100
469/469 - 1s - loss: 0.6236 - sparse_categorical_accuracy: 0.7764 - val_loss: 0.6367 - val_sparse_categorical_accuracy: 0.7744 - 770ms/epoch - 2ms/step

Epoch 30/100
469/469 - 1s - loss: 0.6299 - sparse_categorical_accuracy: 0.7721 - val_loss: 0.6473 - val_sparse_categorical_accuracy: 0.7747 - 829ms/epoch - 2ms/step

Epoch 31/100
469/469 - 1s - loss: 0.6216 - sparse_categorical_accuracy: 0.7749 - val_loss: 0.6450 - val_sparse_categorical_accuracy: 0.7696 - 752ms/epoch - 2ms/step

Epoch 32/100
469/469 - 1s - loss: 0.6188 - sparse_categorical_accuracy: 0.7781 - val_loss: 0.6337 - val_sparse_categorical_accuracy: 0.7793 - 759ms/epoch - 2ms/step

Epoch 33/100
469/469 - 1s - loss: 0.6188 - sparse_categorical_accuracy: 0.7779 - val_loss: 0.6675 - val_sparse_categorical_accuracy: 0.7695 - 777ms/epoch - 2ms/step

Epoch 34/100
469/469 - 1s - loss: 0.6109 - sparse_categorical_accuracy: 0.7798 - val_loss: 0.6541 - val_sparse_categorical_accuracy: 0.7727 - 840ms/epoch - 2ms/step

Epoch 35/100
469/469 - 1s - loss: 0.6141 - sparse_categorical_accuracy: 0.7776 - val_loss: 0.6518 - val_sparse_categorical_accuracy: 0.7776 - 840ms/epoch - 2ms/step

tegorical_accuracy: 0.7694 - 836ms/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.6165 - sparse_categorical_accuracy: 0.7789 - val_loss: 0.6391 - val_sparse_ca
tegorical_accuracy: 0.7716 - 912ms/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.6057 - sparse_categorical_accuracy: 0.7817 - val_loss: 0.6574 - val_sparse_ca
tegorical_accuracy: 0.7703 - 783ms/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.6022 - sparse_categorical_accuracy: 0.7842 - val_loss: 0.6468 - val_sparse_ca
tegorical_accuracy: 0.7612 - 814ms/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.6058 - sparse_categorical_accuracy: 0.7829 - val_loss: 0.6344 - val_sparse_ca
tegorical_accuracy: 0.7790 - 681ms/epoch - 1ms/step
Epoch 40/100
469/469 - 1s - loss: 0.6144 - sparse_categorical_accuracy: 0.7777 - val_loss: 0.6480 - val_sparse_ca
tegorical_accuracy: 0.7710 - 847ms/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.6130 - sparse_categorical_accuracy: 0.7785 - val_loss: 0.7368 - val_sparse_ca
tegorical_accuracy: 0.7411 - 728ms/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.6044 - sparse_categorical_accuracy: 0.7820 - val_loss: 0.6418 - val_sparse_ca
tegorical_accuracy: 0.7759 - 819ms/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.6020 - sparse_categorical_accuracy: 0.7833 - val_loss: 0.6464 - val_sparse_ca
tegorical_accuracy: 0.7716 - 911ms/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.5973 - sparse_categorical_accuracy: 0.7832 - val_loss: 0.6552 - val_sparse_ca
tegorical_accuracy: 0.7685 - 835ms/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.6027 - sparse_categorical_accuracy: 0.7807 - val_loss: 0.6385 - val_sparse_ca
tegorical_accuracy: 0.7708 - 892ms/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.6009 - sparse_categorical_accuracy: 0.7804 - val_loss: 0.6325 - val_sparse_ca
tegorical_accuracy: 0.7780 - 845ms/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.6000 - sparse_categorical_accuracy: 0.7828 - val_loss: 0.7638 - val_sparse_ca
tegorical_accuracy: 0.7476 - 725ms/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.6000 - sparse_categorical_accuracy: 0.7820 - val_loss: 0.6505 - val_sparse_ca
tegorical_accuracy: 0.7619 - 782ms/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.5943 - sparse_categorical_accuracy: 0.7838 - val_loss: 0.6476 - val_sparse_ca
tegorical_accuracy: 0.7577 - 768ms/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.6083 - sparse_categorical_accuracy: 0.7805 - val_loss: 0.6457 - val_sparse_ca
tegorical_accuracy: 0.7663 - 725ms/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.6044 - sparse_categorical_accuracy: 0.7818 - val_loss: 0.6312 - val_sparse_ca
tegorical_accuracy: 0.7784 - 878ms/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.5893 - sparse_categorical_accuracy: 0.7850 - val_loss: 0.6377 - val_sparse_ca
tegorical_accuracy: 0.7714 - 814ms/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.5859 - sparse_categorical_accuracy: 0.7854 - val_loss: 0.6502 - val_sparse_ca
tegorical_accuracy: 0.7667 - 903ms/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.6075 - sparse_categorical_accuracy: 0.7795 - val_loss: 0.6660 - val_sparse_ca
tegorical_accuracy: 0.7588 - 809ms/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.5971 - sparse_categorical_accuracy: 0.7811 - val_loss: 0.6761 - val_sparse_ca
tegorical_accuracy: 0.7614 - 792ms/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.5845 - sparse_categorical_accuracy: 0.7872 - val_loss: 0.6325 - val_sparse_ca
tegorical_accuracy: 0.7772 - 857ms/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 0.5882 - sparse_categorical_accuracy: 0.7873 - val_loss: 0.8063 - val_sparse_ca
tegorical_accuracy: 0.7435 - 864ms/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.5885 - sparse_categorical_accuracy: 0.7850 - val_loss: 0.6462 - val_sparse_ca
tegorical_accuracy: 0.7738 - 859ms/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.6076 - sparse_categorical_accuracy: 0.7782 - val_loss: 0.6241 - val_sparse_ca
tegorical_accuracy: 0.7812 - 774ms/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.5952 - sparse_categorical_accuracy: 0.7835 - val_loss: 0.6291 - val_sparse_ca
tegorical_accuracy: 0.7761 - 815ms/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.5809 - sparse_categorical_accuracy: 0.7879 - val_loss: 0.6236 - val_sparse_ca
tegorical_accuracy: 0.7823 - 763ms/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.5943 - sparse_categorical_accuracy: 0.7833 - val_loss: 0.6349 - val_sparse_ca
tegorical_accuracy: 0.7774 - 750ms/epoch - 2ms/step
Epoch 63/100

469/469 - 1s - loss: 0.6145 - sparse_categorical_accuracy: 0.7795 - val_loss: 0.6460 - val_sparse_categorical_accuracy: 0.7580 - 719ms/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.5892 - sparse_categorical_accuracy: 0.7841 - val_loss: 0.6408 - val_sparse_categorical_accuracy: 0.7528 - 828ms/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.5925 - sparse_categorical_accuracy: 0.7828 - val_loss: 0.6430 - val_sparse_categorical_accuracy: 0.7721 - 844ms/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.5829 - sparse_categorical_accuracy: 0.7882 - val_loss: 0.6194 - val_sparse_categorical_accuracy: 0.7801 - 733ms/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.5684 - sparse_categorical_accuracy: 0.7911 - val_loss: 0.6041 - val_sparse_categorical_accuracy: 0.7865 - 716ms/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 0.5746 - sparse_categorical_accuracy: 0.7906 - val_loss: 0.6283 - val_sparse_categorical_accuracy: 0.7815 - 847ms/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 0.5819 - sparse_categorical_accuracy: 0.7881 - val_loss: 0.6121 - val_sparse_categorical_accuracy: 0.7869 - 739ms/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.5692 - sparse_categorical_accuracy: 0.7921 - val_loss: 0.6323 - val_sparse_categorical_accuracy: 0.7794 - 740ms/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.5740 - sparse_categorical_accuracy: 0.7928 - val_loss: 0.6231 - val_sparse_categorical_accuracy: 0.7865 - 867ms/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.5581 - sparse_categorical_accuracy: 0.7959 - val_loss: 0.5977 - val_sparse_categorical_accuracy: 0.7900 - 767ms/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.5709 - sparse_categorical_accuracy: 0.7936 - val_loss: 0.6450 - val_sparse_categorical_accuracy: 0.7701 - 732ms/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.5650 - sparse_categorical_accuracy: 0.7946 - val_loss: 0.6576 - val_sparse_categorical_accuracy: 0.7693 - 830ms/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.5707 - sparse_categorical_accuracy: 0.7916 - val_loss: 0.6374 - val_sparse_categorical_accuracy: 0.7819 - 835ms/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.5663 - sparse_categorical_accuracy: 0.7949 - val_loss: 0.6192 - val_sparse_categorical_accuracy: 0.7774 - 809ms/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.5621 - sparse_categorical_accuracy: 0.7961 - val_loss: 0.6045 - val_sparse_categorical_accuracy: 0.7907 - 778ms/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.5570 - sparse_categorical_accuracy: 0.7984 - val_loss: 0.6215 - val_sparse_categorical_accuracy: 0.7838 - 932ms/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.5476 - sparse_categorical_accuracy: 0.7990 - val_loss: 0.5925 - val_sparse_categorical_accuracy: 0.7915 - 741ms/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.5513 - sparse_categorical_accuracy: 0.8001 - val_loss: 0.6212 - val_sparse_categorical_accuracy: 0.7856 - 900ms/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.5537 - sparse_categorical_accuracy: 0.7969 - val_loss: 0.6021 - val_sparse_categorical_accuracy: 0.7907 - 854ms/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.5478 - sparse_categorical_accuracy: 0.8020 - val_loss: 0.6628 - val_sparse_categorical_accuracy: 0.7720 - 826ms/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.5699 - sparse_categorical_accuracy: 0.7944 - val_loss: 0.6022 - val_sparse_categorical_accuracy: 0.7938 - 852ms/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.5358 - sparse_categorical_accuracy: 0.8059 - val_loss: 0.5911 - val_sparse_categorical_accuracy: 0.7959 - 876ms/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.5481 - sparse_categorical_accuracy: 0.8006 - val_loss: 0.6038 - val_sparse_categorical_accuracy: 0.7903 - 793ms/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.5409 - sparse_categorical_accuracy: 0.8039 - val_loss: 0.6046 - val_sparse_categorical_accuracy: 0.7966 - 703ms/epoch - 1ms/step
Epoch 87/100
469/469 - 1s - loss: 0.5459 - sparse_categorical_accuracy: 0.8033 - val_loss: 0.6013 - val_sparse_categorical_accuracy: 0.7950 - 802ms/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.5449 - sparse_categorical_accuracy: 0.8027 - val_loss: 0.5885 - val_sparse_categorical_accuracy: 0.7984 - 705ms/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.5546 - sparse_categorical_accuracy: 0.7993 - val_loss: 0.6282 - val_sparse_categorical_accuracy: 0.7868 - 728ms/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.5367 - sparse_categorical_accuracy: 0.8064 - val_loss: 0.5917 - val_sparse_categorical_accuracy: 0.7957 - 782ms/epoch - 2ms/step

```
Epoch 91/100
469/469 - 1s - loss: 0.5514 - sparse_categorical_accuracy: 0.8004 - val_loss: 0.6188 - val_sparse_categorical_accuracy: 0.7850 - 734ms/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.5304 - sparse_categorical_accuracy: 0.8090 - val_loss: 0.6131 - val_sparse_categorical_accuracy: 0.7927 - 823ms/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.5478 - sparse_categorical_accuracy: 0.8001 - val_loss: 0.5942 - val_sparse_categorical_accuracy: 0.7945 - 804ms/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.5375 - sparse_categorical_accuracy: 0.8075 - val_loss: 0.6042 - val_sparse_categorical_accuracy: 0.7939 - 735ms/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.5352 - sparse_categorical_accuracy: 0.8070 - val_loss: 0.6021 - val_sparse_categorical_accuracy: 0.7872 - 753ms/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.5456 - sparse_categorical_accuracy: 0.8041 - val_loss: 0.5963 - val_sparse_categorical_accuracy: 0.7995 - 812ms/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.5344 - sparse_categorical_accuracy: 0.8070 - val_loss: 0.6127 - val_sparse_categorical_accuracy: 0.7935 - 722ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.5382 - sparse_categorical_accuracy: 0.8053 - val_loss: 0.5904 - val_sparse_categorical_accuracy: 0.7980 - 742ms/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.5246 - sparse_categorical_accuracy: 0.8098 - val_loss: 0.7122 - val_sparse_categorical_accuracy: 0.7612 - 814ms/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.5371 - sparse_categorical_accuracy: 0.8067 - val_loss: 0.5943 - val_sparse_categorical_accuracy: 0.7941 - 738ms/epoch - 2ms/step
```

In [42]:

```
model_bias3.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.5943 - sparse_categorical_accuracy: 0.7941 - 235ms/epoch - 749us/step
```

Out[42]:

```
[0.5943122506141663, 0.7940999865531921]
```

Model with random_normal bias initializer

In [43]:

```
model_bias4 = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(16, activation='relu', bias_initializer='random_normal'),
    tf.keras.layers.Dense(16, activation='relu', bias_initializer='random_normal'),
    tf.keras.layers.Dense(10, activation='softmax', bias_initializer='random_normal')
])
```

In [44]:

```
model_bias4.compile(optimizer='adam',
                    loss='sparse_categorical_crossentropy',
                    metrics=['sparse_categorical_accuracy'])
```

In [45]:

```
history_bias4 = model_bias4.fit(x_train, y_train,
                                batch_size=128,
                                epochs=100,
                                validation_data=(x_test, y_test),
                                verbose=2)
```

```
Epoch 1/100
469/469 - 1s - loss: 2.5512 - sparse_categorical_accuracy: 0.2122 - val_loss: 1.7842 - val_sparse_categorical_accuracy: 0.3279 - 1s/epoch - 3ms/step
Epoch 2/100
469/469 - 1s - loss: 1.5387 - sparse_categorical_accuracy: 0.4065 - val_loss: 1.3712 - val_sparse_categorical_accuracy: 0.4846 - 733ms/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 1.2580 - sparse_categorical_accuracy: 0.5124 - val_loss: 1.2646 - val_sparse_categorical_accuracy: 0.5206 - 745ms/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 1.1224 - sparse_categorical_accuracy: 0.5512 - val_loss: 1.1963 - val_sparse_categorical_accuracy: 0.5051 - 720ms/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 1.0531 - sparse_categorical_accuracy: 0.5715 - val_loss: 1.0427 - val_sparse_categorical_accuracy: 0.5206 - 745ms/epoch - 2ms/step
```

tegorical accuracy: 0.5686 - 823ms/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 0.9852 - sparse_categorical_accuracy: 0.5924 - val_loss: 0.9628 - val_sparse_ca
tegorical accuracy: 0.6098 - 862ms/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.9401 - sparse_categorical_accuracy: 0.6084 - val_loss: 0.9658 - val_sparse_ca
tegorical accuracy: 0.6122 - 757ms/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 0.9215 - sparse_categorical_accuracy: 0.6161 - val_loss: 0.9943 - val_sparse_ca
tegorical accuracy: 0.5865 - 783ms/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 0.9102 - sparse_categorical_accuracy: 0.6183 - val_loss: 1.0542 - val_sparse_ca
tegorical accuracy: 0.5767 - 703ms/epoch - 1ms/step
Epoch 10/100
469/469 - 1s - loss: 0.9060 - sparse_categorical_accuracy: 0.6179 - val_loss: 0.9123 - val_sparse_ca
tegorical accuracy: 0.6254 - 740ms/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 0.8902 - sparse_categorical_accuracy: 0.6215 - val_loss: 0.9095 - val_sparse_ca
tegorical accuracy: 0.6210 - 799ms/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.8854 - sparse_categorical_accuracy: 0.6241 - val_loss: 0.9220 - val_sparse_ca
tegorical accuracy: 0.6128 - 780ms/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.8755 - sparse_categorical_accuracy: 0.6281 - val_loss: 0.8818 - val_sparse_ca
tegorical accuracy: 0.6309 - 737ms/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.8741 - sparse_categorical_accuracy: 0.6277 - val_loss: 0.9046 - val_sparse_ca
tegorical accuracy: 0.6155 - 816ms/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 0.8786 - sparse_categorical_accuracy: 0.6276 - val_loss: 0.9096 - val_sparse_ca
tegorical accuracy: 0.6254 - 836ms/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.8591 - sparse_categorical_accuracy: 0.6332 - val_loss: 0.8789 - val_sparse_ca
tegorical accuracy: 0.6318 - 739ms/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.8533 - sparse_categorical_accuracy: 0.6342 - val_loss: 0.8827 - val_sparse_ca
tegorical accuracy: 0.6288 - 788ms/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 0.8499 - sparse_categorical_accuracy: 0.6353 - val_loss: 0.8603 - val_sparse_ca
tegorical accuracy: 0.6380 - 754ms/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.8463 - sparse_categorical_accuracy: 0.6372 - val_loss: 0.9099 - val_sparse_ca
tegorical accuracy: 0.6250 - 713ms/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.8442 - sparse_categorical_accuracy: 0.6357 - val_loss: 0.8751 - val_sparse_ca
tegorical accuracy: 0.6322 - 783ms/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.8436 - sparse_categorical_accuracy: 0.6393 - val_loss: 0.9632 - val_sparse_ca
tegorical accuracy: 0.6341 - 781ms/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.8328 - sparse_categorical_accuracy: 0.6472 - val_loss: 0.8385 - val_sparse_ca
tegorical accuracy: 0.6704 - 790ms/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.7813 - sparse_categorical_accuracy: 0.6884 - val_loss: 0.8053 - val_sparse_ca
tegorical accuracy: 0.6856 - 709ms/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.7559 - sparse_categorical_accuracy: 0.6960 - val_loss: 0.8021 - val_sparse_ca
tegorical accuracy: 0.6750 - 736ms/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.7408 - sparse_categorical_accuracy: 0.7036 - val_loss: 0.8124 - val_sparse_ca
tegorical accuracy: 0.6786 - 800ms/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.7349 - sparse_categorical_accuracy: 0.7026 - val_loss: 0.7814 - val_sparse_ca
tegorical accuracy: 0.6882 - 709ms/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.7210 - sparse_categorical_accuracy: 0.7091 - val_loss: 0.7508 - val_sparse_ca
tegorical accuracy: 0.7154 - 810ms/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.7071 - sparse_categorical_accuracy: 0.7133 - val_loss: 0.7384 - val_sparse_ca
tegorical accuracy: 0.7145 - 797ms/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.6974 - sparse_categorical_accuracy: 0.7167 - val_loss: 0.7115 - val_sparse_ca
tegorical accuracy: 0.7209 - 738ms/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.6844 - sparse_categorical_accuracy: 0.7224 - val_loss: 0.7026 - val_sparse_ca
tegorical accuracy: 0.7181 - 772ms/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.6758 - sparse_categorical_accuracy: 0.7258 - val_loss: 0.6918 - val_sparse_ca
tegorical accuracy: 0.7239 - 787ms/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.6646 - sparse_categorical_accuracy: 0.7281 - val_loss: 0.7363 - val_sparse_ca
tegorical accuracy: 0.7117 - 736ms/epoch - 2ms/step
Epoch 33/100

469/469 - 1s - loss: 0.6611 - sparse_categorical_accuracy: 0.7280 - val_loss: 0.7093 - val_sparse_categorical_accuracy: 0.7131 - 809ms/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.6547 - sparse_categorical_accuracy: 0.7309 - val_loss: 0.6957 - val_sparse_categorical_accuracy: 0.7252 - 739ms/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.6463 - sparse_categorical_accuracy: 0.7341 - val_loss: 0.7519 - val_sparse_categorical_accuracy: 0.7138 - 804ms/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.6423 - sparse_categorical_accuracy: 0.7343 - val_loss: 0.7135 - val_sparse_categorical_accuracy: 0.7195 - 801ms/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.6330 - sparse_categorical_accuracy: 0.7372 - val_loss: 0.7097 - val_sparse_categorical_accuracy: 0.7330 - 800ms/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.6375 - sparse_categorical_accuracy: 0.7365 - val_loss: 0.6825 - val_sparse_categorical_accuracy: 0.7257 - 803ms/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.6313 - sparse_categorical_accuracy: 0.7382 - val_loss: 0.6829 - val_sparse_categorical_accuracy: 0.7309 - 781ms/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 0.6271 - sparse_categorical_accuracy: 0.7391 - val_loss: 0.7109 - val_sparse_categorical_accuracy: 0.7156 - 766ms/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.6287 - sparse_categorical_accuracy: 0.7375 - val_loss: 0.7188 - val_sparse_categorical_accuracy: 0.7210 - 819ms/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.6253 - sparse_categorical_accuracy: 0.7415 - val_loss: 0.6715 - val_sparse_categorical_accuracy: 0.7250 - 736ms/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.6197 - sparse_categorical_accuracy: 0.7435 - val_loss: 0.6797 - val_sparse_categorical_accuracy: 0.7246 - 811ms/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.6157 - sparse_categorical_accuracy: 0.7416 - val_loss: 0.6730 - val_sparse_categorical_accuracy: 0.7315 - 715ms/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.6137 - sparse_categorical_accuracy: 0.7412 - val_loss: 0.6761 - val_sparse_categorical_accuracy: 0.7325 - 745ms/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.6145 - sparse_categorical_accuracy: 0.7412 - val_loss: 0.7524 - val_sparse_categorical_accuracy: 0.7046 - 714ms/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.6166 - sparse_categorical_accuracy: 0.7413 - val_loss: 0.7025 - val_sparse_categorical_accuracy: 0.7187 - 728ms/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.6114 - sparse_categorical_accuracy: 0.7432 - val_loss: 0.6811 - val_sparse_categorical_accuracy: 0.7318 - 713ms/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.6130 - sparse_categorical_accuracy: 0.7430 - val_loss: 0.7027 - val_sparse_categorical_accuracy: 0.7300 - 802ms/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.6031 - sparse_categorical_accuracy: 0.7464 - val_loss: 0.6709 - val_sparse_categorical_accuracy: 0.7342 - 832ms/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.6051 - sparse_categorical_accuracy: 0.7448 - val_loss: 0.6947 - val_sparse_categorical_accuracy: 0.7297 - 786ms/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.6040 - sparse_categorical_accuracy: 0.7449 - val_loss: 0.6836 - val_sparse_categorical_accuracy: 0.7179 - 811ms/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.6059 - sparse_categorical_accuracy: 0.7444 - val_loss: 0.6862 - val_sparse_categorical_accuracy: 0.7286 - 771ms/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.6005 - sparse_categorical_accuracy: 0.7466 - val_loss: 0.6689 - val_sparse_categorical_accuracy: 0.7224 - 789ms/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.5978 - sparse_categorical_accuracy: 0.7462 - val_loss: 0.6670 - val_sparse_categorical_accuracy: 0.7349 - 758ms/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.6011 - sparse_categorical_accuracy: 0.7447 - val_loss: 0.6628 - val_sparse_categorical_accuracy: 0.7329 - 840ms/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 0.5916 - sparse_categorical_accuracy: 0.7494 - val_loss: 0.6608 - val_sparse_categorical_accuracy: 0.7388 - 795ms/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.5962 - sparse_categorical_accuracy: 0.7462 - val_loss: 0.6554 - val_sparse_categorical_accuracy: 0.7394 - 801ms/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.5941 - sparse_categorical_accuracy: 0.7468 - val_loss: 0.6569 - val_sparse_categorical_accuracy: 0.7366 - 763ms/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.5897 - sparse_categorical_accuracy: 0.7493 - val_loss: 0.6564 - val_sparse_categorical_accuracy: 0.7362 - 826ms/epoch - 2ms/step

Epoch 61/100
469/469 - 1s - loss: 0.5917 - sparse_categorical_accuracy: 0.7476 - val_loss: 0.6652 - val_sparse_categorical_accuracy: 0.7370 - 800ms/epoch - 2ms/step

Epoch 62/100
469/469 - 1s - loss: 0.5890 - sparse_categorical_accuracy: 0.7489 - val_loss: 0.6691 - val_sparse_categorical_accuracy: 0.7335 - 784ms/epoch - 2ms/step

Epoch 63/100
469/469 - 1s - loss: 0.5871 - sparse_categorical_accuracy: 0.7497 - val_loss: 0.6678 - val_sparse_categorical_accuracy: 0.7357 - 793ms/epoch - 2ms/step

Epoch 64/100
469/469 - 1s - loss: 0.5823 - sparse_categorical_accuracy: 0.7514 - val_loss: 0.6516 - val_sparse_categorical_accuracy: 0.7385 - 786ms/epoch - 2ms/step

Epoch 65/100
469/469 - 1s - loss: 0.5812 - sparse_categorical_accuracy: 0.7518 - val_loss: 0.6789 - val_sparse_categorical_accuracy: 0.7362 - 731ms/epoch - 2ms/step

Epoch 66/100
469/469 - 1s - loss: 0.5787 - sparse_categorical_accuracy: 0.7521 - val_loss: 0.6887 - val_sparse_categorical_accuracy: 0.7248 - 763ms/epoch - 2ms/step

Epoch 67/100
469/469 - 1s - loss: 0.5833 - sparse_categorical_accuracy: 0.7524 - val_loss: 0.6549 - val_sparse_categorical_accuracy: 0.7393 - 812ms/epoch - 2ms/step

Epoch 68/100
469/469 - 1s - loss: 0.5833 - sparse_categorical_accuracy: 0.7509 - val_loss: 0.6492 - val_sparse_categorical_accuracy: 0.7410 - 789ms/epoch - 2ms/step

Epoch 69/100
469/469 - 1s - loss: 0.5826 - sparse_categorical_accuracy: 0.7554 - val_loss: 0.6801 - val_sparse_categorical_accuracy: 0.7282 - 824ms/epoch - 2ms/step

Epoch 70/100
469/469 - 1s - loss: 0.5783 - sparse_categorical_accuracy: 0.7582 - val_loss: 0.6759 - val_sparse_categorical_accuracy: 0.7479 - 767ms/epoch - 2ms/step

Epoch 71/100
469/469 - 1s - loss: 0.5775 - sparse_categorical_accuracy: 0.7614 - val_loss: 0.6550 - val_sparse_categorical_accuracy: 0.7413 - 735ms/epoch - 2ms/step

Epoch 72/100
469/469 - 1s - loss: 0.5671 - sparse_categorical_accuracy: 0.7720 - val_loss: 0.6564 - val_sparse_categorical_accuracy: 0.7512 - 825ms/epoch - 2ms/step

Epoch 73/100
469/469 - 1s - loss: 0.5653 - sparse_categorical_accuracy: 0.7754 - val_loss: 0.6675 - val_sparse_categorical_accuracy: 0.7583 - 712ms/epoch - 2ms/step

Epoch 74/100
469/469 - 1s - loss: 0.5632 - sparse_categorical_accuracy: 0.7771 - val_loss: 0.6601 - val_sparse_categorical_accuracy: 0.7544 - 739ms/epoch - 2ms/step

Epoch 75/100
469/469 - 1s - loss: 0.5591 - sparse_categorical_accuracy: 0.7781 - val_loss: 0.6433 - val_sparse_categorical_accuracy: 0.7539 - 791ms/epoch - 2ms/step

Epoch 76/100
469/469 - 1s - loss: 0.5559 - sparse_categorical_accuracy: 0.7800 - val_loss: 0.6493 - val_sparse_categorical_accuracy: 0.7642 - 734ms/epoch - 2ms/step

Epoch 77/100
469/469 - 1s - loss: 0.5549 - sparse_categorical_accuracy: 0.7815 - val_loss: 0.6375 - val_sparse_categorical_accuracy: 0.7636 - 747ms/epoch - 2ms/step

Epoch 78/100
469/469 - 1s - loss: 0.5527 - sparse_categorical_accuracy: 0.7831 - val_loss: 0.7143 - val_sparse_categorical_accuracy: 0.7430 - 753ms/epoch - 2ms/step

Epoch 79/100
469/469 - 1s - loss: 0.5605 - sparse_categorical_accuracy: 0.7798 - val_loss: 0.6768 - val_sparse_categorical_accuracy: 0.7510 - 818ms/epoch - 2ms/step

Epoch 80/100
469/469 - 1s - loss: 0.5499 - sparse_categorical_accuracy: 0.7839 - val_loss: 0.6576 - val_sparse_categorical_accuracy: 0.7618 - 808ms/epoch - 2ms/step

Epoch 81/100
469/469 - 1s - loss: 0.5517 - sparse_categorical_accuracy: 0.7836 - val_loss: 0.6405 - val_sparse_categorical_accuracy: 0.7572 - 742ms/epoch - 2ms/step

Epoch 82/100
469/469 - 1s - loss: 0.5486 - sparse_categorical_accuracy: 0.7840 - val_loss: 0.6293 - val_sparse_categorical_accuracy: 0.7667 - 791ms/epoch - 2ms/step

Epoch 83/100
469/469 - 1s - loss: 0.5493 - sparse_categorical_accuracy: 0.7845 - val_loss: 0.6266 - val_sparse_categorical_accuracy: 0.7684 - 826ms/epoch - 2ms/step

Epoch 84/100
469/469 - 1s - loss: 0.5485 - sparse_categorical_accuracy: 0.7867 - val_loss: 0.6434 - val_sparse_categorical_accuracy: 0.7653 - 833ms/epoch - 2ms/step

Epoch 85/100
469/469 - 1s - loss: 0.5452 - sparse_categorical_accuracy: 0.7874 - val_loss: 0.6255 - val_sparse_categorical_accuracy: 0.7660 - 811ms/epoch - 2ms/step

Epoch 86/100
469/469 - 1s - loss: 0.5414 - sparse_categorical_accuracy: 0.7892 - val_loss: 0.6267 - val_sparse_categorical_accuracy: 0.7629 - 817ms/epoch - 2ms/step

Epoch 87/100
469/469 - 1s - loss: 0.5426 - sparse_categorical_accuracy: 0.7901 - val_loss: 0.6355 - val_sparse_categorical_accuracy: 0.7715 - 782ms/epoch - 2ms/step

Epoch 88/100
469/469 - 1s - loss: 0.5445 - sparse_categorical_accuracy: 0.7893 - val_loss: 0.6359 - val_sparse_categorical_accuracy: 0.7715 - 782ms/epoch - 2ms/step

```

tegorical_accuracy: 0.7676 - 825ms/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.5414 - sparse_categorical_accuracy: 0.7901 - val_loss: 0.6315 - val_sparse_ca
tegorical_accuracy: 0.7641 - 745ms/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.5386 - sparse_categorical_accuracy: 0.7926 - val_loss: 0.6402 - val_sparse_ca
tegorical_accuracy: 0.7769 - 737ms/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.5450 - sparse_categorical_accuracy: 0.7926 - val_loss: 0.6339 - val_sparse_ca
tegorical_accuracy: 0.7681 - 835ms/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.5396 - sparse_categorical_accuracy: 0.7949 - val_loss: 0.6296 - val_sparse_ca
tegorical_accuracy: 0.7711 - 745ms/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.5393 - sparse_categorical_accuracy: 0.7927 - val_loss: 0.6574 - val_sparse_ca
tegorical_accuracy: 0.7455 - 842ms/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.5402 - sparse_categorical_accuracy: 0.7926 - val_loss: 0.6405 - val_sparse_ca
tegorical_accuracy: 0.7711 - 800ms/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.5365 - sparse_categorical_accuracy: 0.7946 - val_loss: 0.6639 - val_sparse_ca
tegorical_accuracy: 0.7593 - 834ms/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.5348 - sparse_categorical_accuracy: 0.7975 - val_loss: 0.6321 - val_sparse_ca
tegorical_accuracy: 0.7709 - 750ms/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.5337 - sparse_categorical_accuracy: 0.7960 - val_loss: 0.6232 - val_sparse_ca
tegorical_accuracy: 0.7824 - 778ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.5332 - sparse_categorical_accuracy: 0.7983 - val_loss: 0.6700 - val_sparse_ca
tegorical_accuracy: 0.7709 - 738ms/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.5347 - sparse_categorical_accuracy: 0.7993 - val_loss: 0.6198 - val_sparse_ca
tegorical_accuracy: 0.7792 - 697ms/epoch - 1ms/step
Epoch 100/100
469/469 - 1s - loss: 0.5319 - sparse_categorical_accuracy: 0.8020 - val_loss: 0.6201 - val_sparse_ca
tegorical_accuracy: 0.7849 - 806ms/epoch - 2ms/step

```

In [46]:

```
model_bias4.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.6201 - sparse_categorical_accuracy: 0.7849 - 248ms/epoch - 793us/step
```

Out[46]:

```
[0.6200987100601196, 0.7849000096321106]
```

Model with he_uniform bias initializer

In [47]:

```

model_bias5 = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(16, activation='relu', bias_initializer='he_uniform'),
    tf.keras.layers.Dense(16, activation='relu', bias_initializer='he_uniform'),
    tf.keras.layers.Dense(10, activation='softmax', bias_initializer='he_uniform')
])

```

In [48]:

```

model_bias5.compile(optimizer='adam',
                    loss='sparse_categorical_crossentropy',
                    metrics=['sparse_categorical_accuracy'])

```

In [49]:

```

history_bias5 = model_bias5.fit(x_train, y_train,
                                batch_size=128,
                                epochs=100,
                                validation_data=(x_test, y_test),
                                verbose=2)

```

```

Epoch 1/100
469/469 - 1s - loss: 4.5677 - sparse_categorical_accuracy: 0.1909 - val_loss: 1.9868 - val_sparse_ca
tegorical_accuracy: 0.2593 - 1s/epoch - 3ms/step
Epoch 2/100
469/469 - 1s - loss: 1.8219 - sparse_categorical_accuracy: 0.2916 - val_loss: 1.6956 - val_sparse_ca
tegorical_accuracy: 0.3274 - 822ms/epoch - 2ms/step
Epoch 3/100

```

469/469 - 1s - loss: 1.5646 - sparse_categorical_accuracy: 0.3893 - val_loss: 1.4057 - val_sparse_categorical_accuracy: 0.4558 - 894ms/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 1.2562 - sparse_categorical_accuracy: 0.4864 - val_loss: 1.1929 - val_sparse_categorical_accuracy: 0.4997 - 812ms/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 1.1136 - sparse_categorical_accuracy: 0.5322 - val_loss: 1.0965 - val_sparse_categorical_accuracy: 0.5537 - 742ms/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 1.0683 - sparse_categorical_accuracy: 0.5501 - val_loss: 1.0823 - val_sparse_categorical_accuracy: 0.5506 - 738ms/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 1.0456 - sparse_categorical_accuracy: 0.5560 - val_loss: 1.0809 - val_sparse_categorical_accuracy: 0.5475 - 744ms/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 1.0362 - sparse_categorical_accuracy: 0.5598 - val_loss: 1.0422 - val_sparse_categorical_accuracy: 0.5556 - 836ms/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 1.0180 - sparse_categorical_accuracy: 0.5675 - val_loss: 1.0706 - val_sparse_categorical_accuracy: 0.5658 - 849ms/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 1.0191 - sparse_categorical_accuracy: 0.5703 - val_loss: 1.0237 - val_sparse_categorical_accuracy: 0.5756 - 744ms/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 1.0119 - sparse_categorical_accuracy: 0.5787 - val_loss: 1.0033 - val_sparse_categorical_accuracy: 0.6188 - 759ms/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.9054 - sparse_categorical_accuracy: 0.6421 - val_loss: 0.9487 - val_sparse_categorical_accuracy: 0.6307 - 746ms/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.8597 - sparse_categorical_accuracy: 0.6531 - val_loss: 0.8763 - val_sparse_categorical_accuracy: 0.6611 - 806ms/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.8503 - sparse_categorical_accuracy: 0.6584 - val_loss: 0.8596 - val_sparse_categorical_accuracy: 0.6483 - 796ms/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 0.8421 - sparse_categorical_accuracy: 0.6609 - val_loss: 0.9176 - val_sparse_categorical_accuracy: 0.6585 - 722ms/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.8288 - sparse_categorical_accuracy: 0.6636 - val_loss: 0.8782 - val_sparse_categorical_accuracy: 0.6608 - 742ms/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.8134 - sparse_categorical_accuracy: 0.6704 - val_loss: 0.8773 - val_sparse_categorical_accuracy: 0.6608 - 801ms/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 0.8126 - sparse_categorical_accuracy: 0.6736 - val_loss: 0.8391 - val_sparse_categorical_accuracy: 0.6692 - 787ms/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.8111 - sparse_categorical_accuracy: 0.6719 - val_loss: 0.8252 - val_sparse_categorical_accuracy: 0.6774 - 740ms/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.8017 - sparse_categorical_accuracy: 0.6746 - val_loss: 0.8268 - val_sparse_categorical_accuracy: 0.6723 - 768ms/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.7906 - sparse_categorical_accuracy: 0.6759 - val_loss: 0.8210 - val_sparse_categorical_accuracy: 0.6803 - 738ms/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.7910 - sparse_categorical_accuracy: 0.6780 - val_loss: 0.8377 - val_sparse_categorical_accuracy: 0.6742 - 734ms/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.7842 - sparse_categorical_accuracy: 0.6827 - val_loss: 0.8230 - val_sparse_categorical_accuracy: 0.6797 - 844ms/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.7774 - sparse_categorical_accuracy: 0.6847 - val_loss: 0.8164 - val_sparse_categorical_accuracy: 0.6735 - 710ms/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.7747 - sparse_categorical_accuracy: 0.6835 - val_loss: 0.8535 - val_sparse_categorical_accuracy: 0.6632 - 810ms/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.7851 - sparse_categorical_accuracy: 0.6805 - val_loss: 0.7984 - val_sparse_categorical_accuracy: 0.6876 - 833ms/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.7731 - sparse_categorical_accuracy: 0.6884 - val_loss: 0.8011 - val_sparse_categorical_accuracy: 0.6861 - 703ms/epoch - 1ms/step
Epoch 28/100
469/469 - 1s - loss: 0.7642 - sparse_categorical_accuracy: 0.6899 - val_loss: 0.8194 - val_sparse_categorical_accuracy: 0.6787 - 798ms/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.7651 - sparse_categorical_accuracy: 0.6862 - val_loss: 0.7935 - val_sparse_categorical_accuracy: 0.6779 - 803ms/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.7609 - sparse_categorical_accuracy: 0.6889 - val_loss: 0.8154 - val_sparse_categorical_accuracy: 0.6642 - 805ms/epoch - 2ms/step

Epoch 31/100
469/469 - 1s - loss: 0.7566 - sparse_categorical_accuracy: 0.6866 - val_loss: 0.7790 - val_sparse_categorical_accuracy: 0.6963 - 746ms/epoch - 2ms/step

Epoch 32/100
469/469 - 1s - loss: 0.7574 - sparse_categorical_accuracy: 0.6881 - val_loss: 0.7956 - val_sparse_categorical_accuracy: 0.6937 - 788ms/epoch - 2ms/step

Epoch 33/100
469/469 - 1s - loss: 0.7524 - sparse_categorical_accuracy: 0.6887 - val_loss: 0.8234 - val_sparse_categorical_accuracy: 0.6831 - 712ms/epoch - 2ms/step

Epoch 34/100
469/469 - 1s - loss: 0.7599 - sparse_categorical_accuracy: 0.6904 - val_loss: 0.8164 - val_sparse_categorical_accuracy: 0.6808 - 765ms/epoch - 2ms/step

Epoch 35/100
469/469 - 1s - loss: 0.7586 - sparse_categorical_accuracy: 0.6870 - val_loss: 0.7889 - val_sparse_categorical_accuracy: 0.6946 - 761ms/epoch - 2ms/step

Epoch 36/100
469/469 - 1s - loss: 0.7542 - sparse_categorical_accuracy: 0.6879 - val_loss: 0.8379 - val_sparse_categorical_accuracy: 0.6760 - 812ms/epoch - 2ms/step

Epoch 37/100
469/469 - 1s - loss: 0.7540 - sparse_categorical_accuracy: 0.6865 - val_loss: 0.7893 - val_sparse_categorical_accuracy: 0.6799 - 763ms/epoch - 2ms/step

Epoch 38/100
469/469 - 1s - loss: 0.7359 - sparse_categorical_accuracy: 0.6960 - val_loss: 0.7766 - val_sparse_categorical_accuracy: 0.6902 - 803ms/epoch - 2ms/step

Epoch 39/100
469/469 - 1s - loss: 0.7345 - sparse_categorical_accuracy: 0.6957 - val_loss: 0.7972 - val_sparse_categorical_accuracy: 0.6852 - 730ms/epoch - 2ms/step

Epoch 40/100
469/469 - 1s - loss: 0.7562 - sparse_categorical_accuracy: 0.6895 - val_loss: 0.8322 - val_sparse_categorical_accuracy: 0.6713 - 772ms/epoch - 2ms/step

Epoch 41/100
469/469 - 1s - loss: 0.7539 - sparse_categorical_accuracy: 0.6888 - val_loss: 0.7991 - val_sparse_categorical_accuracy: 0.6854 - 800ms/epoch - 2ms/step

Epoch 42/100
469/469 - 1s - loss: 0.7384 - sparse_categorical_accuracy: 0.6922 - val_loss: 0.8279 - val_sparse_categorical_accuracy: 0.6607 - 775ms/epoch - 2ms/step

Epoch 43/100
469/469 - 1s - loss: 0.7648 - sparse_categorical_accuracy: 0.6809 - val_loss: 0.7903 - val_sparse_categorical_accuracy: 0.6814 - 842ms/epoch - 2ms/step

Epoch 44/100
469/469 - 1s - loss: 0.7512 - sparse_categorical_accuracy: 0.6892 - val_loss: 0.7814 - val_sparse_categorical_accuracy: 0.6889 - 795ms/epoch - 2ms/step

Epoch 45/100
469/469 - 1s - loss: 0.7407 - sparse_categorical_accuracy: 0.6892 - val_loss: 0.7697 - val_sparse_categorical_accuracy: 0.6938 - 771ms/epoch - 2ms/step

Epoch 46/100
469/469 - 1s - loss: 0.7450 - sparse_categorical_accuracy: 0.6882 - val_loss: 0.7756 - val_sparse_categorical_accuracy: 0.6916 - 725ms/epoch - 2ms/step

Epoch 47/100
469/469 - 1s - loss: 0.7665 - sparse_categorical_accuracy: 0.6938 - val_loss: 0.7643 - val_sparse_categorical_accuracy: 0.6925 - 810ms/epoch - 2ms/step

Epoch 48/100
469/469 - 1s - loss: 0.7315 - sparse_categorical_accuracy: 0.6932 - val_loss: 0.7843 - val_sparse_categorical_accuracy: 0.6814 - 793ms/epoch - 2ms/step

Epoch 49/100
469/469 - 1s - loss: 0.7402 - sparse_categorical_accuracy: 0.6897 - val_loss: 0.7665 - val_sparse_categorical_accuracy: 0.6871 - 775ms/epoch - 2ms/step

Epoch 50/100
469/469 - 1s - loss: 0.7188 - sparse_categorical_accuracy: 0.7013 - val_loss: 0.7809 - val_sparse_categorical_accuracy: 0.6858 - 811ms/epoch - 2ms/step

Epoch 51/100
469/469 - 1s - loss: 0.7469 - sparse_categorical_accuracy: 0.6912 - val_loss: 0.7821 - val_sparse_categorical_accuracy: 0.6935 - 771ms/epoch - 2ms/step

Epoch 52/100
469/469 - 1s - loss: 0.7337 - sparse_categorical_accuracy: 0.6962 - val_loss: 0.7785 - val_sparse_categorical_accuracy: 0.6921 - 809ms/epoch - 2ms/step

Epoch 53/100
469/469 - 1s - loss: 0.7358 - sparse_categorical_accuracy: 0.6938 - val_loss: 0.7811 - val_sparse_categorical_accuracy: 0.6843 - 812ms/epoch - 2ms/step

Epoch 54/100
469/469 - 1s - loss: 0.7297 - sparse_categorical_accuracy: 0.6964 - val_loss: 0.7816 - val_sparse_categorical_accuracy: 0.6890 - 805ms/epoch - 2ms/step

Epoch 55/100
469/469 - 1s - loss: 0.7345 - sparse_categorical_accuracy: 0.6940 - val_loss: 0.8497 - val_sparse_categorical_accuracy: 0.6441 - 743ms/epoch - 2ms/step

Epoch 56/100
469/469 - 1s - loss: 0.7285 - sparse_categorical_accuracy: 0.6976 - val_loss: 0.7865 - val_sparse_categorical_accuracy: 0.6832 - 832ms/epoch - 2ms/step

Epoch 57/100
469/469 - 1s - loss: 0.7357 - sparse_categorical_accuracy: 0.6944 - val_loss: 0.7836 - val_sparse_categorical_accuracy: 0.6905 - 795ms/epoch - 2ms/step

Epoch 58/100
469/469 - 1s - loss: 0.7380 - sparse_categorical_accuracy: 0.6921 - val_loss: 0.7874 - val_sparse_categorical_accuracy: 0.6921 - 795ms/epoch - 2ms/step

tegorical accuracy: 0.6870 - 802ms/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.7414 - sparse_categorical_accuracy: 0.6903 - val_loss: 0.8542 - val_sparse_categorical_accuracy: 0.6485 - 749ms/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.7250 - sparse_categorical_accuracy: 0.6974 - val_loss: 0.7809 - val_sparse_categorical_accuracy: 0.6935 - 827ms/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.7575 - sparse_categorical_accuracy: 0.6834 - val_loss: 0.7871 - val_sparse_categorical_accuracy: 0.6900 - 830ms/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.7573 - sparse_categorical_accuracy: 0.6840 - val_loss: 0.7926 - val_sparse_categorical_accuracy: 0.6788 - 837ms/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 0.7269 - sparse_categorical_accuracy: 0.6973 - val_loss: 0.7766 - val_sparse_categorical_accuracy: 0.6881 - 820ms/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.7259 - sparse_categorical_accuracy: 0.6999 - val_loss: 0.7683 - val_sparse_categorical_accuracy: 0.6983 - 710ms/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.7148 - sparse_categorical_accuracy: 0.7021 - val_loss: 0.8585 - val_sparse_categorical_accuracy: 0.6520 - 788ms/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.7287 - sparse_categorical_accuracy: 0.6984 - val_loss: 0.7776 - val_sparse_categorical_accuracy: 0.6963 - 800ms/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.7078 - sparse_categorical_accuracy: 0.7078 - val_loss: 0.7615 - val_sparse_categorical_accuracy: 0.7058 - 788ms/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 0.7434 - sparse_categorical_accuracy: 0.6981 - val_loss: 0.7736 - val_sparse_categorical_accuracy: 0.7020 - 735ms/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 0.7148 - sparse_categorical_accuracy: 0.7065 - val_loss: 0.7695 - val_sparse_categorical_accuracy: 0.7011 - 801ms/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.7500 - sparse_categorical_accuracy: 0.6929 - val_loss: 0.9791 - val_sparse_categorical_accuracy: 0.5893 - 758ms/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.7705 - sparse_categorical_accuracy: 0.6736 - val_loss: 0.8371 - val_sparse_categorical_accuracy: 0.6606 - 799ms/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.7262 - sparse_categorical_accuracy: 0.6967 - val_loss: 0.8306 - val_sparse_categorical_accuracy: 0.6717 - 734ms/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.7212 - sparse_categorical_accuracy: 0.7021 - val_loss: 0.7891 - val_sparse_categorical_accuracy: 0.7003 - 818ms/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.7238 - sparse_categorical_accuracy: 0.7014 - val_loss: 0.7780 - val_sparse_categorical_accuracy: 0.6958 - 836ms/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.7114 - sparse_categorical_accuracy: 0.7057 - val_loss: 0.7712 - val_sparse_categorical_accuracy: 0.7003 - 809ms/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.7197 - sparse_categorical_accuracy: 0.6988 - val_loss: 0.8171 - val_sparse_categorical_accuracy: 0.6787 - 834ms/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.7245 - sparse_categorical_accuracy: 0.7012 - val_loss: 0.7880 - val_sparse_categorical_accuracy: 0.6898 - 788ms/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.7278 - sparse_categorical_accuracy: 0.7008 - val_loss: 0.8041 - val_sparse_categorical_accuracy: 0.7009 - 764ms/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.7230 - sparse_categorical_accuracy: 0.7032 - val_loss: 0.8024 - val_sparse_categorical_accuracy: 0.6830 - 729ms/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.7257 - sparse_categorical_accuracy: 0.7022 - val_loss: 0.7859 - val_sparse_categorical_accuracy: 0.6998 - 749ms/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.7203 - sparse_categorical_accuracy: 0.7028 - val_loss: 1.0061 - val_sparse_categorical_accuracy: 0.6028 - 793ms/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.7886 - sparse_categorical_accuracy: 0.6629 - val_loss: 0.7971 - val_sparse_categorical_accuracy: 0.6883 - 813ms/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.7426 - sparse_categorical_accuracy: 0.6925 - val_loss: 0.8188 - val_sparse_categorical_accuracy: 0.6893 - 807ms/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.7342 - sparse_categorical_accuracy: 0.6931 - val_loss: 0.8058 - val_sparse_categorical_accuracy: 0.6623 - 766ms/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.7138 - sparse_categorical_accuracy: 0.7041 - val_loss: 0.7630 - val_sparse_categorical_accuracy: 0.7024 - 728ms/epoch - 2ms/step
Epoch 86/100

469/469 - 1s - loss: 0.7197 - sparse_categorical_accuracy: 0.7014 - val_loss: 0.8183 - val_sparse_categorical_accuracy: 0.6813 - 845ms/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.7131 - sparse_categorical_accuracy: 0.7032 - val_loss: 0.7658 - val_sparse_categorical_accuracy: 0.6987 - 765ms/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.7238 - sparse_categorical_accuracy: 0.6972 - val_loss: 0.7627 - val_sparse_categorical_accuracy: 0.7020 - 760ms/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.7105 - sparse_categorical_accuracy: 0.7038 - val_loss: 0.8006 - val_sparse_categorical_accuracy: 0.6714 - 747ms/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.7597 - sparse_categorical_accuracy: 0.6818 - val_loss: 0.7552 - val_sparse_categorical_accuracy: 0.7007 - 740ms/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.7069 - sparse_categorical_accuracy: 0.7084 - val_loss: 0.7647 - val_sparse_categorical_accuracy: 0.6947 - 814ms/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.7122 - sparse_categorical_accuracy: 0.7026 - val_loss: 0.7697 - val_sparse_categorical_accuracy: 0.6987 - 835ms/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.7047 - sparse_categorical_accuracy: 0.7089 - val_loss: 0.7781 - val_sparse_categorical_accuracy: 0.6978 - 834ms/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.7139 - sparse_categorical_accuracy: 0.6995 - val_loss: 0.7658 - val_sparse_categorical_accuracy: 0.7004 - 746ms/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.7488 - sparse_categorical_accuracy: 0.6848 - val_loss: 0.8031 - val_sparse_categorical_accuracy: 0.7050 - 793ms/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.7196 - sparse_categorical_accuracy: 0.7035 - val_loss: 0.8012 - val_sparse_categorical_accuracy: 0.7057 - 733ms/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.7255 - sparse_categorical_accuracy: 0.7017 - val_loss: 0.7676 - val_sparse_categorical_accuracy: 0.7041 - 840ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.7494 - sparse_categorical_accuracy: 0.6902 - val_loss: 0.8081 - val_sparse_categorical_accuracy: 0.6856 - 799ms/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.7144 - sparse_categorical_accuracy: 0.7018 - val_loss: 0.7872 - val_sparse_categorical_accuracy: 0.6910 - 803ms/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.7316 - sparse_categorical_accuracy: 0.6975 - val_loss: 0.7703 - val_sparse_categorical_accuracy: 0.7064 - 750ms/epoch - 2ms/step

In [50]:

```
model_bias5.evaluate(x_test, y_test, verbose=2)
```

313/313 - 0s - loss: 0.7703 - sparse_categorical_accuracy: 0.7064 - 232ms/epoch - 741us/step

Out[50]:

[0.7702702879905701, 0.7063999772071838]

Model with he_normal bias initializer

In [51]:

```
model_bias6 = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(16, activation='relu', bias_initializer='he_normal'),
    tf.keras.layers.Dense(16, activation='relu', bias_initializer='he_normal'),
    tf.keras.layers.Dense(10, activation='softmax', bias_initializer='he_normal')
])
```

In [52]:

```
model_bias6.compile(optimizer='adam',
                    loss='sparse_categorical_crossentropy',
                    metrics=['sparse_categorical_accuracy'])
```

In [53]:

```
history_bias6 = model_bias6.fit(x_train, y_train,
                                batch_size=128,
                                epochs=100,
                                validation_data=(x_test, y_test),
                                verbose=2)
```

Epoch 1/100

469/469 - 1s - loss: 3.6881 - sparse_categorical_accuracy: 0.1008 - val_loss: 2.3215 - val_sparse_categorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 2/100
469/469 - 1s - loss: 2.3094 - sparse_categorical_accuracy: 0.1000 - val_loss: 2.3036 - val_sparse_categorical_accuracy: 0.1001 - 848ms/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 2.3035 - sparse_categorical_accuracy: 0.0993 - val_loss: 2.3028 - val_sparse_categorical_accuracy: 0.1001 - 811ms/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 2.1172 - sparse_categorical_accuracy: 0.1758 - val_loss: 1.9811 - val_sparse_categorical_accuracy: 0.2306 - 765ms/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 1.9146 - sparse_categorical_accuracy: 0.2429 - val_loss: 1.8600 - val_sparse_categorical_accuracy: 0.2709 - 737ms/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 1.7819 - sparse_categorical_accuracy: 0.2684 - val_loss: 1.7379 - val_sparse_categorical_accuracy: 0.2718 - 835ms/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 1.7126 - sparse_categorical_accuracy: 0.2775 - val_loss: 1.7639 - val_sparse_categorical_accuracy: 0.2864 - 720ms/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 1.6572 - sparse_categorical_accuracy: 0.2918 - val_loss: 1.5660 - val_sparse_categorical_accuracy: 0.3012 - 829ms/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 1.5329 - sparse_categorical_accuracy: 0.3094 - val_loss: 1.4870 - val_sparse_categorical_accuracy: 0.3154 - 773ms/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 1.4899 - sparse_categorical_accuracy: 0.3262 - val_loss: 1.5356 - val_sparse_categorical_accuracy: 0.2878 - 792ms/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 1.5010 - sparse_categorical_accuracy: 0.3220 - val_loss: 1.5729 - val_sparse_categorical_accuracy: 0.3172 - 802ms/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 1.4854 - sparse_categorical_accuracy: 0.3292 - val_loss: 1.5769 - val_sparse_categorical_accuracy: 0.2980 - 802ms/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 1.4712 - sparse_categorical_accuracy: 0.3282 - val_loss: 1.6272 - val_sparse_categorical_accuracy: 0.2964 - 807ms/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 1.4644 - sparse_categorical_accuracy: 0.3309 - val_loss: 1.4340 - val_sparse_categorical_accuracy: 0.3498 - 816ms/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 1.4563 - sparse_categorical_accuracy: 0.3341 - val_loss: 1.5132 - val_sparse_categorical_accuracy: 0.3164 - 736ms/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 1.4549 - sparse_categorical_accuracy: 0.3446 - val_loss: 1.4592 - val_sparse_categorical_accuracy: 0.3539 - 815ms/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 1.4440 - sparse_categorical_accuracy: 0.3471 - val_loss: 1.4475 - val_sparse_categorical_accuracy: 0.3453 - 810ms/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 1.4404 - sparse_categorical_accuracy: 0.3377 - val_loss: 1.4500 - val_sparse_categorical_accuracy: 0.3622 - 746ms/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 1.4421 - sparse_categorical_accuracy: 0.3441 - val_loss: 1.4403 - val_sparse_categorical_accuracy: 0.3155 - 722ms/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 1.4456 - sparse_categorical_accuracy: 0.3408 - val_loss: 1.4392 - val_sparse_categorical_accuracy: 0.3555 - 804ms/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 1.4341 - sparse_categorical_accuracy: 0.3405 - val_loss: 1.4307 - val_sparse_categorical_accuracy: 0.3414 - 801ms/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 1.4589 - sparse_categorical_accuracy: 0.3380 - val_loss: 1.5362 - val_sparse_categorical_accuracy: 0.3146 - 766ms/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 1.4644 - sparse_categorical_accuracy: 0.3365 - val_loss: 1.4503 - val_sparse_categorical_accuracy: 0.3389 - 817ms/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 1.4247 - sparse_categorical_accuracy: 0.3447 - val_loss: 1.4158 - val_sparse_categorical_accuracy: 0.3585 - 765ms/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 1.4479 - sparse_categorical_accuracy: 0.3359 - val_loss: 1.4301 - val_sparse_categorical_accuracy: 0.3445 - 818ms/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 1.4262 - sparse_categorical_accuracy: 0.3453 - val_loss: 1.4340 - val_sparse_categorical_accuracy: 0.3308 - 826ms/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 1.4303 - sparse_categorical_accuracy: 0.3419 - val_loss: 1.4158 - val_sparse_categorical_accuracy: 0.3657 - 821ms/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 1.4613 - sparse_categorical_accuracy: 0.3371 - val_loss: 1.4204 - val_sparse_categorical_accuracy: 0.3355 - 811ms/epoch - 2ms/step

Epoch 29/100
469/469 - 1s - loss: 1.4380 - sparse_categorical_accuracy: 0.3404 - val_loss: 1.5096 - val_sparse_categorical_accuracy: 0.3310 - 733ms/epoch - 2ms/step

Epoch 30/100
469/469 - 1s - loss: 1.4356 - sparse_categorical_accuracy: 0.3343 - val_loss: 1.4295 - val_sparse_categorical_accuracy: 0.3499 - 805ms/epoch - 2ms/step

Epoch 31/100
469/469 - 1s - loss: 1.4543 - sparse_categorical_accuracy: 0.3275 - val_loss: 1.4301 - val_sparse_categorical_accuracy: 0.3404 - 705ms/epoch - 2ms/step

Epoch 32/100
469/469 - 1s - loss: 1.4241 - sparse_categorical_accuracy: 0.3408 - val_loss: 1.4802 - val_sparse_categorical_accuracy: 0.3286 - 740ms/epoch - 2ms/step

Epoch 33/100
469/469 - 1s - loss: 1.4426 - sparse_categorical_accuracy: 0.3362 - val_loss: 1.4227 - val_sparse_categorical_accuracy: 0.3409 - 727ms/epoch - 2ms/step

Epoch 34/100
469/469 - 1s - loss: 1.4244 - sparse_categorical_accuracy: 0.3456 - val_loss: 1.5119 - val_sparse_categorical_accuracy: 0.3046 - 799ms/epoch - 2ms/step

Epoch 35/100
469/469 - 1s - loss: 1.4419 - sparse_categorical_accuracy: 0.3385 - val_loss: 1.4545 - val_sparse_categorical_accuracy: 0.3711 - 775ms/epoch - 2ms/step

Epoch 36/100
469/469 - 1s - loss: 1.4319 - sparse_categorical_accuracy: 0.3404 - val_loss: 1.4449 - val_sparse_categorical_accuracy: 0.3630 - 806ms/epoch - 2ms/step

Epoch 37/100
469/469 - 1s - loss: 1.4187 - sparse_categorical_accuracy: 0.3485 - val_loss: 1.4655 - val_sparse_categorical_accuracy: 0.3489 - 802ms/epoch - 2ms/step

Epoch 38/100
469/469 - 1s - loss: 1.4525 - sparse_categorical_accuracy: 0.3415 - val_loss: 1.4427 - val_sparse_categorical_accuracy: 0.3545 - 836ms/epoch - 2ms/step

Epoch 39/100
469/469 - 1s - loss: 1.4692 - sparse_categorical_accuracy: 0.3277 - val_loss: 1.4517 - val_sparse_categorical_accuracy: 0.3307 - 714ms/epoch - 2ms/step

Epoch 40/100
469/469 - 1s - loss: 1.4234 - sparse_categorical_accuracy: 0.3454 - val_loss: 1.4870 - val_sparse_categorical_accuracy: 0.3145 - 712ms/epoch - 2ms/step

Epoch 41/100
469/469 - 1s - loss: 1.4301 - sparse_categorical_accuracy: 0.3445 - val_loss: 1.4498 - val_sparse_categorical_accuracy: 0.3550 - 794ms/epoch - 2ms/step

Epoch 42/100
469/469 - 1s - loss: 1.4312 - sparse_categorical_accuracy: 0.3408 - val_loss: 1.4455 - val_sparse_categorical_accuracy: 0.3250 - 845ms/epoch - 2ms/step

Epoch 43/100
469/469 - 1s - loss: 1.4410 - sparse_categorical_accuracy: 0.3421 - val_loss: 1.4191 - val_sparse_categorical_accuracy: 0.3453 - 811ms/epoch - 2ms/step

Epoch 44/100
469/469 - 1s - loss: 1.4455 - sparse_categorical_accuracy: 0.3413 - val_loss: 1.4248 - val_sparse_categorical_accuracy: 0.3447 - 728ms/epoch - 2ms/step

Epoch 45/100
469/469 - 1s - loss: 1.4407 - sparse_categorical_accuracy: 0.3419 - val_loss: 1.5861 - val_sparse_categorical_accuracy: 0.3112 - 748ms/epoch - 2ms/step

Epoch 46/100
469/469 - 1s - loss: 1.4319 - sparse_categorical_accuracy: 0.3439 - val_loss: 1.4498 - val_sparse_categorical_accuracy: 0.3294 - 831ms/epoch - 2ms/step

Epoch 47/100
469/469 - 1s - loss: 1.4251 - sparse_categorical_accuracy: 0.3449 - val_loss: 1.4513 - val_sparse_categorical_accuracy: 0.3331 - 746ms/epoch - 2ms/step

Epoch 48/100
469/469 - 1s - loss: 1.4489 - sparse_categorical_accuracy: 0.3348 - val_loss: 1.4268 - val_sparse_categorical_accuracy: 0.3408 - 747ms/epoch - 2ms/step

Epoch 49/100
469/469 - 1s - loss: 1.4420 - sparse_categorical_accuracy: 0.3424 - val_loss: 1.4258 - val_sparse_categorical_accuracy: 0.3648 - 789ms/epoch - 2ms/step

Epoch 50/100
469/469 - 1s - loss: 1.4429 - sparse_categorical_accuracy: 0.3393 - val_loss: 1.5525 - val_sparse_categorical_accuracy: 0.2991 - 791ms/epoch - 2ms/step

Epoch 51/100
469/469 - 1s - loss: 1.4205 - sparse_categorical_accuracy: 0.3519 - val_loss: 1.5646 - val_sparse_categorical_accuracy: 0.3044 - 793ms/epoch - 2ms/step

Epoch 52/100
469/469 - 1s - loss: 1.4367 - sparse_categorical_accuracy: 0.3488 - val_loss: 1.5563 - val_sparse_categorical_accuracy: 0.3097 - 764ms/epoch - 2ms/step

Epoch 53/100
469/469 - 1s - loss: 1.4224 - sparse_categorical_accuracy: 0.3478 - val_loss: 1.4179 - val_sparse_categorical_accuracy: 0.3586 - 711ms/epoch - 2ms/step

Epoch 54/100
469/469 - 1s - loss: 1.4172 - sparse_categorical_accuracy: 0.3483 - val_loss: 1.4410 - val_sparse_categorical_accuracy: 0.3527 - 774ms/epoch - 2ms/step

Epoch 55/100
469/469 - 1s - loss: 1.4762 - sparse_categorical_accuracy: 0.3311 - val_loss: 1.4845 - val_sparse_categorical_accuracy: 0.3375 - 805ms/epoch - 2ms/step

Epoch 56/100
469/469 - 1s - loss: 1.4266 - sparse_categorical_accuracy: 0.3472 - val_loss: 1.4969 - val_sparse_categorical_accuracy: 0.3472 - 774ms/epoch - 2ms/step

tegorical_accuracy: 0.3566 - 810ms/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 1.4360 - sparse_categorical_accuracy: 0.3447 - val_loss: 1.4007 - val_sparse_ca
tegorical_accuracy: 0.3628 - 791ms/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 1.4205 - sparse_categorical_accuracy: 0.3511 - val_loss: 1.3966 - val_sparse_ca
tegorical_accuracy: 0.3574 - 772ms/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 1.4211 - sparse_categorical_accuracy: 0.3486 - val_loss: 1.5615 - val_sparse_ca
tegorical_accuracy: 0.3231 - 836ms/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 1.4713 - sparse_categorical_accuracy: 0.3334 - val_loss: 1.4457 - val_sparse_ca
tegorical_accuracy: 0.3259 - 784ms/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 1.4176 - sparse_categorical_accuracy: 0.3463 - val_loss: 1.5516 - val_sparse_ca
tegorical_accuracy: 0.3167 - 827ms/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 1.4296 - sparse_categorical_accuracy: 0.3409 - val_loss: 1.4229 - val_sparse_ca
tegorical_accuracy: 0.3461 - 801ms/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 1.4085 - sparse_categorical_accuracy: 0.3518 - val_loss: 1.4417 - val_sparse_ca
tegorical_accuracy: 0.3412 - 787ms/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 1.4472 - sparse_categorical_accuracy: 0.3402 - val_loss: 1.4255 - val_sparse_ca
tegorical_accuracy: 0.3555 - 726ms/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 1.4287 - sparse_categorical_accuracy: 0.3479 - val_loss: 1.4224 - val_sparse_ca
tegorical_accuracy: 0.3529 - 765ms/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 1.4340 - sparse_categorical_accuracy: 0.3446 - val_loss: 1.4334 - val_sparse_ca
tegorical_accuracy: 0.3716 - 812ms/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 1.4499 - sparse_categorical_accuracy: 0.3408 - val_loss: 1.4482 - val_sparse_ca
tegorical_accuracy: 0.3364 - 746ms/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 1.4277 - sparse_categorical_accuracy: 0.3447 - val_loss: 1.4191 - val_sparse_ca
tegorical_accuracy: 0.3402 - 799ms/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 1.4071 - sparse_categorical_accuracy: 0.3528 - val_loss: 1.4314 - val_sparse_ca
tegorical_accuracy: 0.3454 - 749ms/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 1.4362 - sparse_categorical_accuracy: 0.3459 - val_loss: 1.4414 - val_sparse_ca
tegorical_accuracy: 0.3445 - 742ms/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 1.4469 - sparse_categorical_accuracy: 0.3487 - val_loss: 1.4142 - val_sparse_ca
tegorical_accuracy: 0.3503 - 798ms/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 1.4098 - sparse_categorical_accuracy: 0.3579 - val_loss: 1.4163 - val_sparse_ca
tegorical_accuracy: 0.3617 - 784ms/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 1.4137 - sparse_categorical_accuracy: 0.3501 - val_loss: 1.4376 - val_sparse_ca
tegorical_accuracy: 0.3378 - 859ms/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 1.4494 - sparse_categorical_accuracy: 0.3426 - val_loss: 1.4345 - val_sparse_ca
tegorical_accuracy: 0.3465 - 800ms/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 1.4193 - sparse_categorical_accuracy: 0.3523 - val_loss: 1.4093 - val_sparse_ca
tegorical_accuracy: 0.3614 - 774ms/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 1.4455 - sparse_categorical_accuracy: 0.3446 - val_loss: 1.5019 - val_sparse_ca
tegorical_accuracy: 0.3261 - 766ms/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 1.4524 - sparse_categorical_accuracy: 0.3422 - val_loss: 1.4352 - val_sparse_ca
tegorical_accuracy: 0.3384 - 772ms/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 1.4394 - sparse_categorical_accuracy: 0.3535 - val_loss: 1.4251 - val_sparse_ca
tegorical_accuracy: 0.3439 - 716ms/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 1.4026 - sparse_categorical_accuracy: 0.3564 - val_loss: 1.4162 - val_sparse_ca
tegorical_accuracy: 0.3668 - 795ms/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 1.4052 - sparse_categorical_accuracy: 0.3510 - val_loss: 1.4178 - val_sparse_ca
tegorical_accuracy: 0.3683 - 745ms/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 1.4107 - sparse_categorical_accuracy: 0.3545 - val_loss: 1.4149 - val_sparse_ca
tegorical_accuracy: 0.3653 - 768ms/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 1.4174 - sparse_categorical_accuracy: 0.3545 - val_loss: 1.4325 - val_sparse_ca
tegorical_accuracy: 0.3381 - 730ms/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 1.4120 - sparse_categorical_accuracy: 0.3539 - val_loss: 1.4318 - val_sparse_ca
tegorical_accuracy: 0.3377 - 733ms/epoch - 2ms/step
Epoch 84/100

```

469/469 - 1s - loss: 1.4594 - sparse_categorical_accuracy: 0.3409 - val_loss: 1.5002 - val_sparse_categorical_accuracy: 0.3238 - 814ms/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 1.4304 - sparse_categorical_accuracy: 0.3470 - val_loss: 1.4128 - val_sparse_categorical_accuracy: 0.3542 - 798ms/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 1.4874 - sparse_categorical_accuracy: 0.3313 - val_loss: 1.4183 - val_sparse_categorical_accuracy: 0.3637 - 769ms/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 1.4001 - sparse_categorical_accuracy: 0.3516 - val_loss: 1.4067 - val_sparse_categorical_accuracy: 0.3617 - 805ms/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 1.4171 - sparse_categorical_accuracy: 0.3494 - val_loss: 1.4345 - val_sparse_categorical_accuracy: 0.3253 - 933ms/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 1.4182 - sparse_categorical_accuracy: 0.3480 - val_loss: 1.4125 - val_sparse_categorical_accuracy: 0.3309 - 817ms/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 1.4490 - sparse_categorical_accuracy: 0.3391 - val_loss: 1.4505 - val_sparse_categorical_accuracy: 0.3407 - 775ms/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 1.4436 - sparse_categorical_accuracy: 0.3455 - val_loss: 1.5148 - val_sparse_categorical_accuracy: 0.3211 - 781ms/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 1.4375 - sparse_categorical_accuracy: 0.3444 - val_loss: 1.4240 - val_sparse_categorical_accuracy: 0.3356 - 743ms/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 1.4192 - sparse_categorical_accuracy: 0.3532 - val_loss: 1.4397 - val_sparse_categorical_accuracy: 0.3364 - 815ms/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 1.4063 - sparse_categorical_accuracy: 0.3622 - val_loss: 1.4466 - val_sparse_categorical_accuracy: 0.3674 - 744ms/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 1.4401 - sparse_categorical_accuracy: 0.3431 - val_loss: 1.4442 - val_sparse_categorical_accuracy: 0.3297 - 798ms/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 1.4583 - sparse_categorical_accuracy: 0.3399 - val_loss: 1.4648 - val_sparse_categorical_accuracy: 0.3450 - 775ms/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 1.4178 - sparse_categorical_accuracy: 0.3486 - val_loss: 1.4589 - val_sparse_categorical_accuracy: 0.3415 - 796ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 1.4409 - sparse_categorical_accuracy: 0.3403 - val_loss: 1.4395 - val_sparse_categorical_accuracy: 0.3694 - 777ms/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 1.4483 - sparse_categorical_accuracy: 0.3451 - val_loss: 1.5575 - val_sparse_categorical_accuracy: 0.3395 - 783ms/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 1.4587 - sparse_categorical_accuracy: 0.3330 - val_loss: 1.4661 - val_sparse_categorical_accuracy: 0.3383 - 767ms/epoch - 2ms/step

```

In [54]:

```
model_bias6.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 1.4661 - sparse_categorical_accuracy: 0.3383 - 239ms/epoch - 762us/step
```

Out[54]:

```
[1.4660979509353638, 0.3382999897003174]
```

Best initializers in these choices

Model with kernel_initializer='random_normal' or 'random_uniform', bias_initializer='zeros' or 'random_normal'

In [55]:

```

model_init = tf.keras.Sequential([
    tf.keras.layers.Flatten(input_shape=(28, 28)),
    tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_normal', bias_initializer='random_normal'),
    tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_normal', bias_initializer='random_normal'),
    tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_normal', bias_initializer='random_normal')
])

```

In [56]:

```
model_init.compile(optimizer='adam',  
                    loss='sparse_categorical_crossentropy',  
                    metrics=['sparse_categorical_accuracy'])
```

In [57]:

```
history_init = model_init.fit(x_train, y_train,  
                              batch_size=128,  
                              epochs=100,  
                              validation_data=(x_test, y_test),  
                              verbose=2  
                              )
```

```
Epoch 1/100  
469/469 - 1s - loss: 0.7004 - sparse_categorical_accuracy: 0.7493 - val_loss: 0.5460 - val_sparse_ca  
tegorical_accuracy: 0.7973 - 1s/epoch - 2ms/step  
Epoch 2/100  
469/469 - 1s - loss: 0.4747 - sparse_categorical_accuracy: 0.8295 - val_loss: 0.4748 - val_sparse_ca  
tegorical_accuracy: 0.8286 - 869ms/epoch - 2ms/step  
Epoch 3/100  
469/469 - 1s - loss: 0.4261 - sparse_categorical_accuracy: 0.8478 - val_loss: 0.4928 - val_sparse_ca  
tegorical_accuracy: 0.8260 - 769ms/epoch - 2ms/step  
Epoch 4/100  
469/469 - 1s - loss: 0.4069 - sparse_categorical_accuracy: 0.8531 - val_loss: 0.4469 - val_sparse_ca  
tegorical_accuracy: 0.8433 - 846ms/epoch - 2ms/step  
Epoch 5/100  
469/469 - 1s - loss: 0.3893 - sparse_categorical_accuracy: 0.8596 - val_loss: 0.4405 - val_sparse_ca  
tegorical_accuracy: 0.8441 - 722ms/epoch - 2ms/step  
Epoch 6/100  
469/469 - 1s - loss: 0.3779 - sparse_categorical_accuracy: 0.8644 - val_loss: 0.4399 - val_sparse_ca  
tegorical_accuracy: 0.8423 - 748ms/epoch - 2ms/step  
Epoch 7/100  
469/469 - 1s - loss: 0.3654 - sparse_categorical_accuracy: 0.8664 - val_loss: 0.4313 - val_sparse_ca  
tegorical_accuracy: 0.8453 - 844ms/epoch - 2ms/step  
Epoch 8/100  
469/469 - 1s - loss: 0.3582 - sparse_categorical_accuracy: 0.8709 - val_loss: 0.4180 - val_sparse_ca  
tegorical_accuracy: 0.8551 - 820ms/epoch - 2ms/step  
Epoch 9/100  
469/469 - 1s - loss: 0.3515 - sparse_categorical_accuracy: 0.8724 - val_loss: 0.4248 - val_sparse_ca  
tegorical_accuracy: 0.8501 - 821ms/epoch - 2ms/step  
Epoch 10/100  
469/469 - 1s - loss: 0.3467 - sparse_categorical_accuracy: 0.8724 - val_loss: 0.4037 - val_sparse_ca  
tegorical_accuracy: 0.8571 - 766ms/epoch - 2ms/step  
Epoch 11/100  
469/469 - 1s - loss: 0.3420 - sparse_categorical_accuracy: 0.8741 - val_loss: 0.4167 - val_sparse_ca  
tegorical_accuracy: 0.8557 - 809ms/epoch - 2ms/step  
Epoch 12/100  
469/469 - 1s - loss: 0.3365 - sparse_categorical_accuracy: 0.8766 - val_loss: 0.4474 - val_sparse_ca  
tegorical_accuracy: 0.8389 - 724ms/epoch - 2ms/step  
Epoch 13/100  
469/469 - 1s - loss: 0.3356 - sparse_categorical_accuracy: 0.8775 - val_loss: 0.4154 - val_sparse_ca  
tegorical_accuracy: 0.8563 - 827ms/epoch - 2ms/step  
Epoch 14/100  
469/469 - 1s - loss: 0.3292 - sparse_categorical_accuracy: 0.8796 - val_loss: 0.4207 - val_sparse_ca  
tegorical_accuracy: 0.8595 - 821ms/epoch - 2ms/step  
Epoch 15/100  
469/469 - 1s - loss: 0.3317 - sparse_categorical_accuracy: 0.8779 - val_loss: 0.4045 - val_sparse_ca  
tegorical_accuracy: 0.8592 - 832ms/epoch - 2ms/step  
Epoch 16/100  
469/469 - 1s - loss: 0.3254 - sparse_categorical_accuracy: 0.8804 - val_loss: 0.4181 - val_sparse_ca  
tegorical_accuracy: 0.8521 - 834ms/epoch - 2ms/step  
Epoch 17/100  
469/469 - 1s - loss: 0.3236 - sparse_categorical_accuracy: 0.8817 - val_loss: 0.4019 - val_sparse_ca  
tegorical_accuracy: 0.8631 - 864ms/epoch - 2ms/step  
Epoch 18/100  
469/469 - 1s - loss: 0.3173 - sparse_categorical_accuracy: 0.8844 - val_loss: 0.4101 - val_sparse_ca  
tegorical_accuracy: 0.8582 - 853ms/epoch - 2ms/step  
Epoch 19/100  
469/469 - 1s - loss: 0.3203 - sparse_categorical_accuracy: 0.8836 - val_loss: 0.4050 - val_sparse_ca  
tegorical_accuracy: 0.8624 - 854ms/epoch - 2ms/step  
Epoch 20/100  
469/469 - 1s - loss: 0.3145 - sparse_categorical_accuracy: 0.8847 - val_loss: 0.4194 - val_sparse_ca  
tegorical_accuracy: 0.8566 - 808ms/epoch - 2ms/step  
Epoch 21/100  
469/469 - 1s - loss: 0.3117 - sparse_categorical_accuracy: 0.8854 - val_loss: 0.3995 - val_sparse_ca  
tegorical_accuracy: 0.8620 - 829ms/epoch - 2ms/step  
Epoch 22/100  
469/469 - 1s - loss: 0.3111 - sparse_categorical_accuracy: 0.8852 - val_loss: 0.4122 - val_sparse_ca  
tegorical_accuracy: 0.8591 - 844ms/epoch - 2ms/step  
Epoch 23/100
```

469/469 - 1s - loss: 0.3117 - sparse_categorical_accuracy: 0.8869 - val_loss: 0.4153 - val_sparse_categorical_accuracy: 0.8576 - 860ms/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.3054 - sparse_categorical_accuracy: 0.8874 - val_loss: 0.4136 - val_sparse_categorical_accuracy: 0.8587 - 763ms/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.3068 - sparse_categorical_accuracy: 0.8871 - val_loss: 0.4111 - val_sparse_categorical_accuracy: 0.8589 - 842ms/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.3031 - sparse_categorical_accuracy: 0.8880 - val_loss: 0.4150 - val_sparse_categorical_accuracy: 0.8616 - 788ms/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.3043 - sparse_categorical_accuracy: 0.8889 - val_loss: 0.4153 - val_sparse_categorical_accuracy: 0.8580 - 793ms/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.3013 - sparse_categorical_accuracy: 0.8896 - val_loss: 0.4110 - val_sparse_categorical_accuracy: 0.8612 - 756ms/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.2978 - sparse_categorical_accuracy: 0.8900 - val_loss: 0.4216 - val_sparse_categorical_accuracy: 0.8544 - 753ms/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.3042 - sparse_categorical_accuracy: 0.8892 - val_loss: 0.4217 - val_sparse_categorical_accuracy: 0.8578 - 815ms/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.2962 - sparse_categorical_accuracy: 0.8919 - val_loss: 0.4327 - val_sparse_categorical_accuracy: 0.8536 - 753ms/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.3002 - sparse_categorical_accuracy: 0.8909 - val_loss: 0.4043 - val_sparse_categorical_accuracy: 0.8635 - 760ms/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.2930 - sparse_categorical_accuracy: 0.8919 - val_loss: 0.4290 - val_sparse_categorical_accuracy: 0.8599 - 824ms/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.2988 - sparse_categorical_accuracy: 0.8898 - val_loss: 0.5203 - val_sparse_categorical_accuracy: 0.8383 - 725ms/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.2957 - sparse_categorical_accuracy: 0.8924 - val_loss: 0.4159 - val_sparse_categorical_accuracy: 0.8573 - 832ms/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.2914 - sparse_categorical_accuracy: 0.8923 - val_loss: 0.4246 - val_sparse_categorical_accuracy: 0.8601 - 806ms/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.2928 - sparse_categorical_accuracy: 0.8919 - val_loss: 0.4327 - val_sparse_categorical_accuracy: 0.8588 - 948ms/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.2924 - sparse_categorical_accuracy: 0.8922 - val_loss: 0.4217 - val_sparse_categorical_accuracy: 0.8591 - 985ms/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.2917 - sparse_categorical_accuracy: 0.8925 - val_loss: 0.4304 - val_sparse_categorical_accuracy: 0.8577 - 960ms/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 0.2950 - sparse_categorical_accuracy: 0.8921 - val_loss: 0.4434 - val_sparse_categorical_accuracy: 0.8590 - 856ms/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.2905 - sparse_categorical_accuracy: 0.8927 - val_loss: 0.4224 - val_sparse_categorical_accuracy: 0.8577 - 958ms/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.2869 - sparse_categorical_accuracy: 0.8946 - val_loss: 0.4355 - val_sparse_categorical_accuracy: 0.8532 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.2891 - sparse_categorical_accuracy: 0.8939 - val_loss: 0.4271 - val_sparse_categorical_accuracy: 0.8593 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.2848 - sparse_categorical_accuracy: 0.8938 - val_loss: 0.4512 - val_sparse_categorical_accuracy: 0.8545 - 761ms/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.2849 - sparse_categorical_accuracy: 0.8948 - val_loss: 0.4233 - val_sparse_categorical_accuracy: 0.8658 - 835ms/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.2875 - sparse_categorical_accuracy: 0.8938 - val_loss: 0.4320 - val_sparse_categorical_accuracy: 0.8595 - 746ms/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.2823 - sparse_categorical_accuracy: 0.8963 - val_loss: 0.4335 - val_sparse_categorical_accuracy: 0.8631 - 750ms/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.2825 - sparse_categorical_accuracy: 0.8963 - val_loss: 0.4386 - val_sparse_categorical_accuracy: 0.8599 - 845ms/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.2805 - sparse_categorical_accuracy: 0.8964 - val_loss: 0.4432 - val_sparse_categorical_accuracy: 0.8609 - 797ms/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.2790 - sparse_categorical_accuracy: 0.8982 - val_loss: 0.4352 - val_sparse_categorical_accuracy: 0.8598 - 813ms/epoch - 2ms/step

Epoch 51/100
469/469 - 1s - loss: 0.2899 - sparse_categorical_accuracy: 0.8946 - val_loss: 0.4587 - val_sparse_categorical_accuracy: 0.8544 - 850ms/epoch - 2ms/step

Epoch 52/100
469/469 - 1s - loss: 0.2822 - sparse_categorical_accuracy: 0.8951 - val_loss: 0.4358 - val_sparse_categorical_accuracy: 0.8612 - 843ms/epoch - 2ms/step

Epoch 53/100
469/469 - 1s - loss: 0.2804 - sparse_categorical_accuracy: 0.8978 - val_loss: 0.4351 - val_sparse_categorical_accuracy: 0.8577 - 839ms/epoch - 2ms/step

Epoch 54/100
469/469 - 1s - loss: 0.2751 - sparse_categorical_accuracy: 0.8982 - val_loss: 0.4728 - val_sparse_categorical_accuracy: 0.8508 - 757ms/epoch - 2ms/step

Epoch 55/100
469/469 - 1s - loss: 0.2788 - sparse_categorical_accuracy: 0.8973 - val_loss: 0.4533 - val_sparse_categorical_accuracy: 0.8559 - 755ms/epoch - 2ms/step

Epoch 56/100
469/469 - 1s - loss: 0.2787 - sparse_categorical_accuracy: 0.8961 - val_loss: 0.4508 - val_sparse_categorical_accuracy: 0.8546 - 769ms/epoch - 2ms/step

Epoch 57/100
469/469 - 1s - loss: 0.2816 - sparse_categorical_accuracy: 0.8966 - val_loss: 0.4554 - val_sparse_categorical_accuracy: 0.8570 - 855ms/epoch - 2ms/step

Epoch 58/100
469/469 - 1s - loss: 0.2731 - sparse_categorical_accuracy: 0.8992 - val_loss: 0.4693 - val_sparse_categorical_accuracy: 0.8496 - 847ms/epoch - 2ms/step

Epoch 59/100
469/469 - 1s - loss: 0.2854 - sparse_categorical_accuracy: 0.8957 - val_loss: 0.4473 - val_sparse_categorical_accuracy: 0.8578 - 732ms/epoch - 2ms/step

Epoch 60/100
469/469 - 1s - loss: 0.2793 - sparse_categorical_accuracy: 0.8970 - val_loss: 0.4612 - val_sparse_categorical_accuracy: 0.8561 - 780ms/epoch - 2ms/step

Epoch 61/100
469/469 - 1s - loss: 0.2719 - sparse_categorical_accuracy: 0.8996 - val_loss: 0.4658 - val_sparse_categorical_accuracy: 0.8548 - 797ms/epoch - 2ms/step

Epoch 62/100
469/469 - 1s - loss: 0.2759 - sparse_categorical_accuracy: 0.8975 - val_loss: 0.4497 - val_sparse_categorical_accuracy: 0.8552 - 802ms/epoch - 2ms/step

Epoch 63/100
469/469 - 1s - loss: 0.2728 - sparse_categorical_accuracy: 0.8989 - val_loss: 0.4405 - val_sparse_categorical_accuracy: 0.8636 - 864ms/epoch - 2ms/step

Epoch 64/100
469/469 - 1s - loss: 0.2754 - sparse_categorical_accuracy: 0.8988 - val_loss: 0.4425 - val_sparse_categorical_accuracy: 0.8650 - 732ms/epoch - 2ms/step

Epoch 65/100
469/469 - 1s - loss: 0.2723 - sparse_categorical_accuracy: 0.8997 - val_loss: 0.4316 - val_sparse_categorical_accuracy: 0.8639 - 817ms/epoch - 2ms/step

Epoch 66/100
469/469 - 1s - loss: 0.2770 - sparse_categorical_accuracy: 0.8974 - val_loss: 0.4672 - val_sparse_categorical_accuracy: 0.8611 - 738ms/epoch - 2ms/step

Epoch 67/100
469/469 - 1s - loss: 0.2708 - sparse_categorical_accuracy: 0.8994 - val_loss: 0.4771 - val_sparse_categorical_accuracy: 0.8494 - 825ms/epoch - 2ms/step

Epoch 68/100
469/469 - 1s - loss: 0.2718 - sparse_categorical_accuracy: 0.9000 - val_loss: 0.4953 - val_sparse_categorical_accuracy: 0.8519 - 744ms/epoch - 2ms/step

Epoch 69/100
469/469 - 1s - loss: 0.2697 - sparse_categorical_accuracy: 0.8999 - val_loss: 0.4632 - val_sparse_categorical_accuracy: 0.8525 - 786ms/epoch - 2ms/step

Epoch 70/100
469/469 - 1s - loss: 0.2727 - sparse_categorical_accuracy: 0.8986 - val_loss: 0.4761 - val_sparse_categorical_accuracy: 0.8548 - 834ms/epoch - 2ms/step

Epoch 71/100
469/469 - 1s - loss: 0.2711 - sparse_categorical_accuracy: 0.9002 - val_loss: 0.4657 - val_sparse_categorical_accuracy: 0.8585 - 777ms/epoch - 2ms/step

Epoch 72/100
469/469 - 1s - loss: 0.2724 - sparse_categorical_accuracy: 0.8993 - val_loss: 0.4597 - val_sparse_categorical_accuracy: 0.8549 - 757ms/epoch - 2ms/step

Epoch 73/100
469/469 - 1s - loss: 0.2707 - sparse_categorical_accuracy: 0.8995 - val_loss: 0.4454 - val_sparse_categorical_accuracy: 0.8629 - 723ms/epoch - 2ms/step

Epoch 74/100
469/469 - 1s - loss: 0.2686 - sparse_categorical_accuracy: 0.9015 - val_loss: 0.4726 - val_sparse_categorical_accuracy: 0.8543 - 766ms/epoch - 2ms/step

Epoch 75/100
469/469 - 1s - loss: 0.2669 - sparse_categorical_accuracy: 0.9011 - val_loss: 0.4719 - val_sparse_categorical_accuracy: 0.8573 - 796ms/epoch - 2ms/step

Epoch 76/100
469/469 - 1s - loss: 0.2753 - sparse_categorical_accuracy: 0.8988 - val_loss: 0.4668 - val_sparse_categorical_accuracy: 0.8586 - 770ms/epoch - 2ms/step

Epoch 77/100
469/469 - 1s - loss: 0.2691 - sparse_categorical_accuracy: 0.9013 - val_loss: 0.4570 - val_sparse_categorical_accuracy: 0.8616 - 816ms/epoch - 2ms/step

Epoch 78/100
469/469 - 1s - loss: 0.2707 - sparse_categorical_accuracy: 0.9010 - val_loss: 0.4662 - val_sparse_categorical_accuracy: 0.8616 - 816ms/epoch - 2ms/step

tegorical accuracy: 0.8594 - 748ms/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.2685 - sparse_categorical_accuracy: 0.9001 - val_loss: 0.4583 - val_sparse_categorical_accuracy: 0.8603 - 808ms/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.2681 - sparse_categorical_accuracy: 0.9003 - val_loss: 0.4672 - val_sparse_categorical_accuracy: 0.8609 - 872ms/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.2677 - sparse_categorical_accuracy: 0.9010 - val_loss: 0.4659 - val_sparse_categorical_accuracy: 0.8594 - 761ms/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.2716 - sparse_categorical_accuracy: 0.8996 - val_loss: 0.4640 - val_sparse_categorical_accuracy: 0.8547 - 746ms/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.2643 - sparse_categorical_accuracy: 0.9024 - val_loss: 0.4642 - val_sparse_categorical_accuracy: 0.8583 - 835ms/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.2660 - sparse_categorical_accuracy: 0.9015 - val_loss: 0.4602 - val_sparse_categorical_accuracy: 0.8591 - 773ms/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.2655 - sparse_categorical_accuracy: 0.9017 - val_loss: 0.4667 - val_sparse_categorical_accuracy: 0.8578 - 780ms/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.2685 - sparse_categorical_accuracy: 0.9011 - val_loss: 0.4709 - val_sparse_categorical_accuracy: 0.8600 - 805ms/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.2655 - sparse_categorical_accuracy: 0.9026 - val_loss: 0.4782 - val_sparse_categorical_accuracy: 0.8556 - 797ms/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.2660 - sparse_categorical_accuracy: 0.9011 - val_loss: 0.4644 - val_sparse_categorical_accuracy: 0.8601 - 771ms/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.2675 - sparse_categorical_accuracy: 0.9013 - val_loss: 0.5005 - val_sparse_categorical_accuracy: 0.8586 - 793ms/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.2641 - sparse_categorical_accuracy: 0.9025 - val_loss: 0.4911 - val_sparse_categorical_accuracy: 0.8563 - 854ms/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.2651 - sparse_categorical_accuracy: 0.9026 - val_loss: 0.4655 - val_sparse_categorical_accuracy: 0.8624 - 746ms/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.2643 - sparse_categorical_accuracy: 0.9011 - val_loss: 0.4740 - val_sparse_categorical_accuracy: 0.8618 - 781ms/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.2620 - sparse_categorical_accuracy: 0.9032 - val_loss: 0.5013 - val_sparse_categorical_accuracy: 0.8520 - 760ms/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.2641 - sparse_categorical_accuracy: 0.9018 - val_loss: 0.4970 - val_sparse_categorical_accuracy: 0.8596 - 813ms/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.2660 - sparse_categorical_accuracy: 0.9016 - val_loss: 0.4714 - val_sparse_categorical_accuracy: 0.8610 - 814ms/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.2620 - sparse_categorical_accuracy: 0.9036 - val_loss: 0.5153 - val_sparse_categorical_accuracy: 0.8533 - 835ms/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.2620 - sparse_categorical_accuracy: 0.9034 - val_loss: 0.4898 - val_sparse_categorical_accuracy: 0.8594 - 853ms/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.2651 - sparse_categorical_accuracy: 0.9028 - val_loss: 0.4754 - val_sparse_categorical_accuracy: 0.8602 - 815ms/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.2657 - sparse_categorical_accuracy: 0.9022 - val_loss: 0.4790 - val_sparse_categorical_accuracy: 0.8589 - 725ms/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.2621 - sparse_categorical_accuracy: 0.9036 - val_loss: 0.4912 - val_sparse_categorical_accuracy: 0.8580 - 821ms/epoch - 2ms/step

In [58]:

```
model_init.evaluate(x_test, y_test, verbose=2)
```

313/313 - 0s - loss: 0.4912 - sparse_categorical_accuracy: 0.8580 - 236ms/epoch - 755us/step

Out[58]:

[0.491211861371994, 0.8579999804496765]

Plots

In [59]:

```
import matplotlib.pyplot as plt
```

Plots for different kernel initializers

In [60]:

```
training_accuracy_kernel0 = history.history['sparse_categorical_accuracy']
validation_accuracy_kernel0 = history.history['val_sparse_categorical_accuracy']

training_accuracy_kernel1 = history_kernel1.history['sparse_categorical_accuracy']
validation_accuracy_kernel1 = history_kernel1.history['val_sparse_categorical_accuracy']

training_accuracy_kernel2 = history_kernel2.history['sparse_categorical_accuracy']
validation_accuracy_kernel2 = history_kernel2.history['val_sparse_categorical_accuracy']

training_accuracy_kernel3 = history_kernel3.history['sparse_categorical_accuracy']
validation_accuracy_kernel3 = history_kernel3.history['val_sparse_categorical_accuracy']

training_accuracy_kernel4 = history_kernel4.history['sparse_categorical_accuracy']
validation_accuracy_kernel4 = history_kernel4.history['val_sparse_categorical_accuracy']

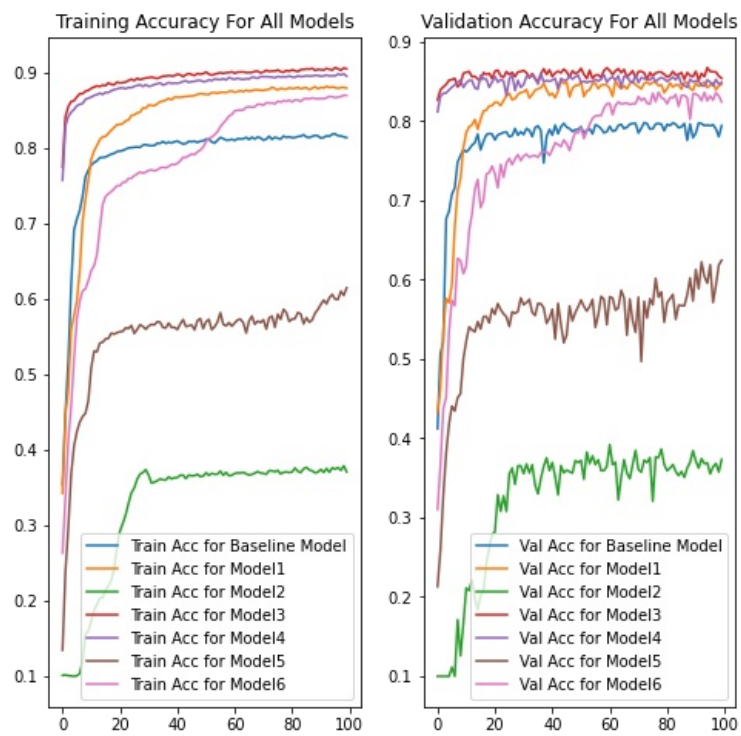
training_accuracy_kernel5 = history_kernel5.history['sparse_categorical_accuracy']
validation_accuracy_kernel5 = history_kernel5.history['val_sparse_categorical_accuracy']

training_accuracy_kernel6 = history_kernel6.history['sparse_categorical_accuracy']
validation_accuracy_kernel6 = history_kernel6.history['val_sparse_categorical_accuracy']

epochs_range=range(100)

plt.figure(figsize=(8, 8))
plt.subplot(1, 2, 1)
plt.plot(epochs_range, training_accuracy_kernel0, label='Train Acc for Baseline Model')
plt.plot(epochs_range, training_accuracy_kernel1, label='Train Acc for Model1')
plt.plot(epochs_range, training_accuracy_kernel2, label='Train Acc for Model2')
plt.plot(epochs_range, training_accuracy_kernel3, label='Train Acc for Model3')
plt.plot(epochs_range, training_accuracy_kernel4, label='Train Acc for Model4')
plt.plot(epochs_range, training_accuracy_kernel5, label='Train Acc for Model5')
plt.plot(epochs_range, training_accuracy_kernel6, label='Train Acc for Model6')
plt.legend(loc='lower right')
plt.title('Training Accuracy For All Models')

plt.subplot(1, 2, 2)
plt.plot(epochs_range, validation_accuracy_kernel0, label='Val Acc for Baseline Model')
plt.plot(epochs_range, validation_accuracy_kernel1, label='Val Acc for Model1')
plt.plot(epochs_range, validation_accuracy_kernel2, label='Val Acc for Model2')
plt.plot(epochs_range, validation_accuracy_kernel3, label='Val Acc for Model3')
plt.plot(epochs_range, validation_accuracy_kernel4, label='Val Acc for Model4')
plt.plot(epochs_range, validation_accuracy_kernel5, label='Val Acc for Model5')
plt.plot(epochs_range, validation_accuracy_kernel6, label='Val Acc for Model6')
plt.legend(loc='lower right')
plt.title('Validation Accuracy For All Models')
plt.show()
```



Plots for different bias initializers

In [61]:

```
training_accuracy_bias0 = history.history['sparse_categorical_accuracy']
validation_accuracy_bias0 = history.history['val_sparse_categorical_accuracy']

training_accuracy_bias1 = history_bias1.history['sparse_categorical_accuracy']
validation_accuracy_bias1 = history_bias1.history['val_sparse_categorical_accuracy']

training_accuracy_bias2 = history_bias2.history['sparse_categorical_accuracy']
validation_accuracy_bias2 = history_bias2.history['val_sparse_categorical_accuracy']

training_accuracy_bias3 = history_bias3.history['sparse_categorical_accuracy']
validation_accuracy_bias3 = history_bias3.history['val_sparse_categorical_accuracy']

training_accuracy_bias4 = history_bias4.history['sparse_categorical_accuracy']
validation_accuracy_bias4 = history_bias4.history['val_sparse_categorical_accuracy']

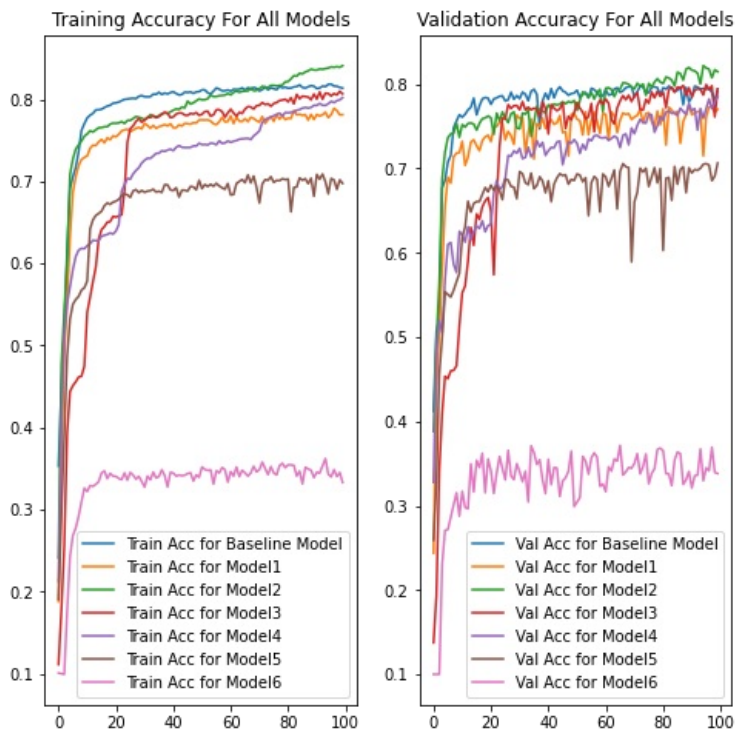
training_accuracy_bias5 = history_bias5.history['sparse_categorical_accuracy']
validation_accuracy_bias5 = history_bias5.history['val_sparse_categorical_accuracy']

training_accuracy_bias6 = history_bias6.history['sparse_categorical_accuracy']
validation_accuracy_bias6 = history_bias6.history['val_sparse_categorical_accuracy']

epochs_range=range(100)

plt.figure(figsize=(8, 8))
plt.subplot(1, 2, 1)
plt.plot(epochs_range, training_accuracy_bias0, label='Train Acc for Baseline Model')
plt.plot(epochs_range, training_accuracy_bias1, label='Train Acc for Model1')
plt.plot(epochs_range, training_accuracy_bias2, label='Train Acc for Model2')
plt.plot(epochs_range, training_accuracy_bias3, label='Train Acc for Model3')
plt.plot(epochs_range, training_accuracy_bias4, label='Train Acc for Model4')
plt.plot(epochs_range, training_accuracy_bias5, label='Train Acc for Model5')
plt.plot(epochs_range, training_accuracy_bias6, label='Train Acc for Model6')
plt.legend(loc='lower right')
plt.title('Training Accuracy For All Models')

plt.subplot(1, 2, 2)
plt.plot(epochs_range, validation_accuracy_bias0, label='Val Acc for Baseline Model')
plt.plot(epochs_range, validation_accuracy_bias1, label='Val Acc for Model1')
plt.plot(epochs_range, validation_accuracy_bias2, label='Val Acc for Model2')
plt.plot(epochs_range, validation_accuracy_bias3, label='Val Acc for Model3')
plt.plot(epochs_range, validation_accuracy_bias4, label='Val Acc for Model4')
plt.plot(epochs_range, validation_accuracy_bias5, label='Val Acc for Model5')
plt.plot(epochs_range, validation_accuracy_bias6, label='Val Acc for Model6')
plt.legend(loc='lower right')
plt.title('Validation Accuracy For All Models')
plt.show()
```



Plots for different initializers

In [62]:

```
training_accuracy0 = history.history['sparse_categorical_accuracy']
validation_accuracy0 = history.history['val_sparse_categorical_accuracy']

training_accuracy_init = history_init.history['sparse_categorical_accuracy']
validation_accuracy_init = history_init.history['val_sparse_categorical_accuracy']

epochs_range=range(100)

plt.figure(figsize=(8, 8))
plt.subplot(1, 2, 1)
plt.plot(epochs_range, training_accuracy0, label='Train Acc for Baseline Model')
plt.plot(epochs_range, training_accuracy_init, label='Train Acc for "best" Model')
plt.legend(loc='lower right')
plt.title('Training Accuracy For All Models')

plt.subplot(1, 2, 2)
plt.plot(epochs_range, validation_accuracy0, label='Val Acc for Baseline Model')
plt.plot(epochs_range, validation_accuracy_init, label='Val Acc for "best" Model1')
plt.legend(loc='lower right')
plt.title('Validation Accuracy For All Models')
plt.show()
```

