In [59]:

```
import tensorflow as tf
import matplotlib.pyplot as plt
```

Baseline model

In [2]:

```
# Load Data
(x_train, y_train), (x_test, y_test) = tf.keras.datasets.mnist.load_data()
```

```
Downloading data from https://storage.googleapis.com/tensorflow/tf-keras-datasets/mnist.npz
11493376/11490434 [==============================] - 1s 0us/step
11501568/11490434 [==============================] - 1s 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 [9]:

```
# 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 - 2s - loss: 3.3832 - sparse_categorical_accuracy: 0.1673 - val_loss: 2.0299 - val_sparse_ca
tegorical_accuracy: 0.2294 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 1.7632 - sparse_categorical_accuracy: 0.3222 - val_loss: 1.5297 - val_sparse_ca
tegorical_accuracy: 0.4045 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 1.4479 - sparse_categorical_accuracy: 0.4098 - val_loss: 1.4049 - val_sparse_ca
tegorical_accuracy: 0.4196 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 1.2899 - sparse_categorical_accuracy: 0.4723 - val_loss: 1.1958 - val_sparse_ca
tegorical_accuracy: 0.5006 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 1.1358 - sparse_categorical_accuracy: 0.5324 - val_loss: 1.0813 - val_sparse_ca
tegorical_accuracy: 0.5769 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 1.0069 - sparse_categorical_accuracy: 0.5863 - val_loss: 0.9630 - val_sparse_ca
tegorical_accuracy: 0.6253 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.9192 - sparse_categorical_accuracy: 0.6399 - val_loss: 0.8979 - val_sparse_ca
tegorical_accuracy: 0.6592 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 0.8308 - sparse_categorical_accuracy: 0.6941 - val_loss: 0.8184 - val_sparse_ca
tegorical_accuracy: 0.7119 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 0.7547 - sparse_categorical_accuracy: 0.7463 - val_loss: 0.7161 - val_sparse_ca
tegorical_accuracy: 0.7768 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 0.6757 - sparse_categorical_accuracy: 0.7950 - val_loss: 0.6302 - val_sparse_ca
tegorical_accuracy: 0.8165 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 0.5868 - sparse_categorical_accuracy: 0.8314 - val_loss: 0.5383 - val_sparse_ca
tegorical_accuracy: 0.8450 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.5180 - sparse_categorical_accuracy: 0.8539 - val_loss: 0.5065 - val_sparse_ca
tegorical_accuracy: 0.8594 - 1s/epoch - 2ms/step
```

```
Epoch 13/100
469/469 - 1s - loss: 0.4885 - sparse_categorical_accuracy: 0.8608 - val_loss: 0.4892 - val_sparse_ca
tegorical_accuracy: 0.8580 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.4662 - sparse_categorical_accuracy: 0.8679 - val_loss: 0.4999 - val_sparse_ca
tegorical_accuracy: 0.8540 - 1s/epoch - 3ms/step
Epoch 15/100
469/469 - 2s - loss: 0.4519 - sparse_categorical_accuracy: 0.8700 - val_loss: 0.4356 - val_sparse_ca
tegorical_accuracy: 0.8774 - 2s/epoch - 4ms/step
Epoch 16/100
469/469 - 1s - loss: 0.4310 - sparse_categorical_accuracy: 0.8774 - val_loss: 0.4296 - val_sparse_ca
tegorical_accuracy: 0.8782 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.4219 - sparse_categorical_accuracy: 0.8784 - val_loss: 0.4354 - val_sparse_ca
tegorical_accuracy: 0.8779 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 0.4121 - sparse_categorical_accuracy: 0.8833 - val_loss: 0.4228 - val_sparse_ca
tegorical_accuracy: 0.8855 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.4020 - sparse_categorical_accuracy: 0.8843 - val_loss: 0.4182 - val_sparse_ca
tegorical_accuracy: 0.8843 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.3946 - sparse_categorical_accuracy: 0.8872 - val_loss: 0.4037 - val_sparse_ca
tegorical_accuracy: 0.8887 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.3852 - sparse_categorical_accuracy: 0.8892 - val_loss: 0.4099 - val_sparse_ca
tegorical_accuracy: 0.8864 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.3824 - sparse_categorical_accuracy: 0.8893 - val_loss: 0.4182 - val_sparse_ca
tegorical_accuracy: 0.8837 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.3734 - sparse_categorical_accuracy: 0.8942 - val_loss: 0.3978 - val_sparse_ca
tegorical_accuracy: 0.8866 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.3680 - sparse_categorical_accuracy: 0.8948 - val_loss: 0.3951 - val_sparse_ca
tegorical_accuracy: 0.8891 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.3638 - sparse_categorical_accuracy: 0.8966 - val_loss: 0.3856 - val_sparse_ca
tegorical_accuracy: 0.8960 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.3563 - sparse_categorical_accuracy: 0.8985 - val_loss: 0.3966 - val_sparse_ca
tegorical_accuracy: 0.8893 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.3468 - sparse_categorical_accuracy: 0.9008 - val_loss: 0.3973 - val_sparse_ca
tegorical_accuracy: 0.8856 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.3384 - sparse_categorical_accuracy: 0.9028 - val_loss: 0.3565 - val_sparse_ca
tegorical_accuracy: 0.8967 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.3352 - sparse_categorical_accuracy: 0.9049 - val_loss: 0.3606 - val_sparse_ca
tegorical_accuracy: 0.9018 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.3317 - sparse_categorical_accuracy: 0.9052 - val_loss: 0.3464 - val_sparse_ca
tegorical_accuracy: 0.9043 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.3272 - sparse_categorical_accuracy: 0.9060 - val_loss: 0.3663 - val_sparse_ca
tegorical_accuracy: 0.8973 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.3203 - sparse_categorical_accuracy: 0.9086 - val_loss: 0.3416 - val_sparse_ca
tegorical_accuracy: 0.9060 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.3181 - sparse_categorical_accuracy: 0.9085 - val_loss: 0.3647 - val_sparse_ca
tegorical_accuracy: 0.8977 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.3190 - sparse_categorical_accuracy: 0.9090 - val_loss: 0.3516 - val_sparse_ca
tegorical_accuracy: 0.9011 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.3155 - sparse_categorical_accuracy: 0.9097 - val_loss: 0.3815 - val_sparse_ca
tegorical_accuracy: 0.8945 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.3129 - sparse_categorical_accuracy: 0.9101 - val_loss: 0.3622 - val_sparse_ca
tegorical_accuracy: 0.8999 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.3100 - sparse_categorical_accuracy: 0.9102 - val_loss: 0.3624 - val_sparse_ca
tegorical_accuracy: 0.8964 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.3089 - sparse_categorical_accuracy: 0.9122 - val_loss: 0.3403 - val_sparse_ca
tegorical_accuracy: 0.9052 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.3059 - sparse_categorical_accuracy: 0.9133 - val_loss: 0.3600 - val_sparse_ca
tegorical_accuracy: 0.8978 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 0.3027 - sparse_categorical_accuracy: 0.9129 - val_loss: 0.3523 - val_sparse_ca
```

```
tegorical_accuracy: 0.9040 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.3008 - sparse_categorical_accuracy: 0.9137 - val_loss: 0.3484 - val_sparse_ca
tegorical_accuracy: 0.9021 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.3005 - sparse_categorical_accuracy: 0.9136 - val_loss: 0.3329 - val_sparse_ca
tegorical_accuracy: 0.9103 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.2955 - sparse_categorical_accuracy: 0.9142 - val_loss: 0.3473 - val_sparse_ca
tegorical_accuracy: 0.9042 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.2948 - sparse_categorical_accuracy: 0.9147 - val_loss: 0.3787 - val_sparse_ca
tegorical_accuracy: 0.9002 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.2951 - sparse_categorical_accuracy: 0.9156 - val_loss: 0.3448 - val_sparse_ca
tegorical_accuracy: 0.9080 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.2947 - sparse_categorical_accuracy: 0.9161 - val_loss: 0.3547 - val_sparse_ca
tegorical_accuracy: 0.9069 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.2927 - sparse_categorical_accuracy: 0.9151 - val_loss: 0.3441 - val_sparse_ca
tegorical_accuracy: 0.9057 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.2941 - sparse_categorical_accuracy: 0.9158 - val_loss: 0.3491 - val_sparse_ca
tegorical_accuracy: 0.9040 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.2911 - sparse_categorical_accuracy: 0.9157 - val_loss: 0.3461 - val_sparse_ca
tegorical_accuracy: 0.9055 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.2875 - sparse_categorical_accuracy: 0.9168 - val_loss: 0.3535 - val_sparse_ca
tegorical_accuracy: 0.9031 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.2874 - sparse_categorical_accuracy: 0.9164 - val_loss: 0.3359 - val_sparse_ca
tegorical_accuracy: 0.9103 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.2887 - sparse_categorical_accuracy: 0.9174 - val_loss: 0.3390 - val_sparse_ca
tegorical_accuracy: 0.9041 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.2857 - sparse_categorical_accuracy: 0.9176 - val_loss: 0.3518 - val_sparse_ca
tegorical_accuracy: 0.9033 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.2865 - sparse_categorical_accuracy: 0.9177 - val_loss: 0.3400 - val_sparse_ca
tegorical_accuracy: 0.9061 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.2837 - sparse_categorical_accuracy: 0.9182 - val_loss: 0.3687 - val_sparse_ca
tegorical_accuracy: 0.8999 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.2840 - sparse_categorical_accuracy: 0.9190 - val_loss: 0.3359 - val_sparse_ca
tegorical_accuracy: 0.9092 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 0.2812 - sparse_categorical_accuracy: 0.9188 - val_loss: 0.3320 - val_sparse_ca
tegorical_accuracy: 0.9115 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.2823 - sparse_categorical_accuracy: 0.9184 - val_loss: 0.3377 - val_sparse_ca
tegorical_accuracy: 0.9064 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.2804 - sparse_categorical_accuracy: 0.9207 - val_loss: 0.3396 - val_sparse_ca
tegorical_accuracy: 0.9075 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.2799 - sparse_categorical_accuracy: 0.9193 - val_loss: 0.3699 - val_sparse_ca
tegorical_accuracy: 0.8952 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.2804 - sparse_categorical_accuracy: 0.9196 - val_loss: 0.3352 - val_sparse_ca
tegorical_accuracy: 0.9090 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.2762 - sparse_categorical_accuracy: 0.9197 - val_loss: 0.3414 - val_sparse_ca
tegorical_accuracy: 0.9045 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 0.2739 - sparse_categorical_accuracy: 0.9204 - val_loss: 0.3378 - val_sparse_ca
tegorical_accuracy: 0.9084 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.2779 - sparse_categorical_accuracy: 0.9196 - val_loss: 0.3462 - val_sparse_ca
tegorical_accuracy: 0.9064 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.2761 - sparse_categorical_accuracy: 0.9205 - val_loss: 0.3210 - val_sparse_ca
tegorical_accuracy: 0.9134 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.2760 - sparse_categorical_accuracy: 0.9204 - val_loss: 0.3259 - val_sparse_ca
tegorical_accuracy: 0.9105 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.2752 - sparse_categorical_accuracy: 0.9205 - val_loss: 0.3434 - val_sparse_ca
tegorical_accuracy: 0.9068 - 1s/epoch - 2ms/step
Epoch 68/100
```

```
469/469 - 1s - loss: 0.2763 - sparse_categorical_accuracy: 0.9205 - val_loss: 0.3252 - val_sparse_ca
tegorical_accuracy: 0.9109 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 0.2756 - sparse_categorical_accuracy: 0.9198 - val_loss: 0.3306 - val_sparse_ca
tegorical_accuracy: 0.9105 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.2746 - sparse_categorical_accuracy: 0.9205 - val_loss: 0.3385 - val_sparse_ca
tegorical_accuracy: 0.9099 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.2722 - sparse_categorical_accuracy: 0.9219 - val_loss: 0.3345 - val_sparse_ca
tegorical_accuracy: 0.9088 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.2703 - sparse_categorical_accuracy: 0.9208 - val_loss: 0.3454 - val_sparse_ca
tegorical_accuracy: 0.9080 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.2691 - sparse_categorical_accuracy: 0.9219 - val_loss: 0.3289 - val_sparse_ca
tegorical_accuracy: 0.9130 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.2706 - sparse_categorical_accuracy: 0.9218 - val_loss: 0.3409 - val_sparse_ca
tegorical_accuracy: 0.9098 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.2698 - sparse_categorical_accuracy: 0.9215 - val_loss: 0.3344 - val_sparse_ca
tegorical_accuracy: 0.9086 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.2671 - sparse_categorical_accuracy: 0.9223 - val_loss: 0.3259 - val_sparse_ca
tegorical_accuracy: 0.9123 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.2664 - sparse_categorical_accuracy: 0.9226 - val_loss: 0.3237 - val_sparse_ca
tegorical_accuracy: 0.9103 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.2681 - sparse_categorical_accuracy: 0.9225 - val_loss: 0.3362 - val_sparse_ca
tegorical_accuracy: 0.9071 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.2659 - sparse_categorical_accuracy: 0.9234 - val_loss: 0.3268 - val_sparse_ca
tegorical_accuracy: 0.9098 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.2641 - sparse_categorical_accuracy: 0.9238 - val_loss: 0.3328 - val_sparse_ca
tegorical_accuracy: 0.9100 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.2664 - sparse_categorical_accuracy: 0.9234 - val_loss: 0.3255 - val_sparse_ca
tegorical_accuracy: 0.9122 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.2648 - sparse_categorical_accuracy: 0.9230 - val_loss: 0.3293 - val_sparse_ca
tegorical_accuracy: 0.9111 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.2647 - sparse_categorical_accuracy: 0.9235 - val_loss: 0.3251 - val_sparse_ca
tegorical_accuracy: 0.9107 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.2638 - sparse_categorical_accuracy: 0.9226 - val_loss: 0.3276 - val_sparse_ca
tegorical_accuracy: 0.9112 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.2643 - sparse_categorical_accuracy: 0.9231 - val_loss: 0.3251 - val_sparse_ca
tegorical_accuracy: 0.9109 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.2644 - sparse_categorical_accuracy: 0.9225 - val_loss: 0.3232 - val_sparse_ca
tegorical_accuracy: 0.9141 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.2652 - sparse_categorical_accuracy: 0.9236 - val_loss: 0.3272 - val_sparse_ca
tegorical_accuracy: 0.9111 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.2638 - sparse_categorical_accuracy: 0.9232 - val_loss: 0.3500 - val_sparse_ca
tegorical_accuracy: 0.9040 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.2625 - sparse_categorical_accuracy: 0.9230 - val_loss: 0.3254 - val_sparse_ca
tegorical_accuracy: 0.9120 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.2608 - sparse_categorical_accuracy: 0.9250 - val_loss: 0.3311 - val_sparse_ca
tegorical_accuracy: 0.9114 - 1s/epoch - 3ms/step
Epoch 91/100
469/469 - 2s - loss: 0.2590 - sparse_categorical_accuracy: 0.9254 - val_loss: 0.3257 - val_sparse_ca
tegorical_accuracy: 0.9114 - 2s/epoch - 4ms/step
Epoch 92/100
469/469 - 2s - loss: 0.2633 - sparse_categorical_accuracy: 0.9231 - val_loss: 0.3239 - val_sparse_ca
tegorical_accuracy: 0.9125 - 2s/epoch - 4ms/step
Epoch 93/100
469/469 - 1s - loss: 0.2600 - sparse_categorical_accuracy: 0.9240 - val_loss: 0.3293 - val_sparse_ca
tegorical_accuracy: 0.9101 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.2592 - sparse_categorical_accuracy: 0.9238 - val_loss: 0.3429 - val_sparse_ca
tegorical_accuracy: 0.9076 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.2614 - sparse_categorical_accuracy: 0.9245 - val_loss: 0.3252 - val_sparse_ca
tegorical_accuracy: 0.9119 - 1s/epoch - 2ms/step
```

```
Epoch 96/100
469/469 - 1s - loss: 0.2629 - sparse_categorical_accuracy: 0.9233 - val_loss: 0.3279 - val_sparse_ca
tegorical_accuracy: 0.9080 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.2612 - sparse_categorical_accuracy: 0.9239 - val_loss: 0.3321 - val_sparse_ca
tegorical_accuracy: 0.9090 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.2605 - sparse_categorical_accuracy: 0.9242 - val_loss: 0.3315 - val_sparse_ca
tegorical_accuracy: 0.9111 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.2590 - sparse_categorical_accuracy: 0.9245 - val_loss: 0.3308 - val_sparse_ca
tegorical_accuracy: 0.9069 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.2591 - sparse_categorical_accuracy: 0.9250 - val_loss: 0.3295 - val_sparse_ca
tegorical_accuracy: 0.9100 - 1s/epoch - 2ms/step
```

In [10]:

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

```
313/313 - 0s - loss: 0.3295 - sparse_categorical_accuracy: 0.9100 - 393ms/epoch - 1ms/step
```

Out[10]:

```
[0.32945308089256287, 0.9100000262260437]
```

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

Rugularization Tuning

Model1: Regularization penalty is L1 with parameter 10^{-1}

In [26]:

```
model1 = tf.keras.models.Sequential([
   tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform',
  activity_regularizer = tf.keras.regularizers.L1(0.1)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
```

In [27]:

```
# Fix the learning rate to be 0.001
model1.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [28]:

```
history1 = model1.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 2s - loss: 2.5008 - sparse_categorical_accuracy: 0.1121 - val_loss: 2.3013 - val_sparse_ca
tegorical_accuracy: 0.1135 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 3ms/step
Epoch 4/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 3ms/step
Epoch 6/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
```

```
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 3ms/step
Epoch 10/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 3ms/step
Epoch 12/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 3ms/step
Epoch 13/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 3ms/step
Epoch 18/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 2s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 2s/epoch - 4ms/step
Epoch 20/100
469/469 - 2s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 2s/epoch - 3ms/step
Epoch 21/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 3ms/step
Epoch 32/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 35/100
```

```
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 3ms/step
Epoch 49/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 3ms/step
Epoch 58/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 3ms/step
```

```
Epoch 63/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 3ms/step
Epoch 81/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 3ms/step
Epoch 84/100
469/469 - 2s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 2s/epoch - 4ms/step
Epoch 85/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
```

```
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
```

In [29]:

```
model1.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 2.3011 - sparse_categorical_accuracy: 0.1135 - 397ms/epoch - 1ms/step
```

Out[29]:

```
[2.3011412620544434, 0.11349999904632568]
```

Model2: Regularization penalty is L1 with parameter 10^{-2}

In [30]:

```
model2 = tf.keras.models.Sequential([
   tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform',
  activity_regularizer = tf.keras.regularizers.L1(0.01)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
```

In [31]:

```
# Fix the learning rate to be 0.001
model2.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [32]:

```
history2 = model2.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 2s - loss: 2.3192 - sparse_categorical_accuracy: 0.1117 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 3ms/step
Epoch 3/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
```

```
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 2.2157 - sparse_categorical_accuracy: 0.1733 - val_loss: 2.1644 - val_sparse_ca
tegorical_accuracy: 0.1836 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 2.1433 - sparse_categorical_accuracy: 0.1936 - val_loss: 2.1204 - val_sparse_ca
tegorical_accuracy: 0.1993 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 2.1263 - sparse_categorical_accuracy: 0.1959 - val_loss: 2.2128 - val_sparse_ca
tegorical_accuracy: 0.2075 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 2.1218 - sparse_categorical_accuracy: 0.1988 - val_loss: 2.1197 - val_sparse_ca
tegorical_accuracy: 0.1997 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 2.0723 - sparse_categorical_accuracy: 0.2218 - val_loss: 2.0178 - val_sparse_ca
tegorical_accuracy: 0.2551 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 2.0172 - sparse_categorical_accuracy: 0.2382 - val_loss: 2.0209 - val_sparse_ca
tegorical_accuracy: 0.2421 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 1.9828 - sparse_categorical_accuracy: 0.2473 - val_loss: 1.9590 - val_sparse_ca
tegorical_accuracy: 0.2704 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 1.9448 - sparse_categorical_accuracy: 0.2521 - val_loss: 1.9342 - val_sparse_ca
tegorical_accuracy: 0.2641 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 1.9192 - sparse_categorical_accuracy: 0.2540 - val_loss: 1.9381 - val_sparse_ca
tegorical_accuracy: 0.2488 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 1.9010 - sparse_categorical_accuracy: 0.2555 - val_loss: 1.9280 - val_sparse_ca
tegorical_accuracy: 0.2459 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 1.9032 - sparse_categorical_accuracy: 0.2537 - val_loss: 1.8898 - val_sparse_ca
tegorical_accuracy: 0.2754 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 1.8903 - sparse_categorical_accuracy: 0.2568 - val_loss: 1.8831 - val_sparse_ca
tegorical_accuracy: 0.2678 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 1.8816 - sparse_categorical_accuracy: 0.2596 - val_loss: 1.8930 - val_sparse_ca
tegorical_accuracy: 0.2671 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 1.8791 - sparse_categorical_accuracy: 0.2603 - val_loss: 1.9290 - val_sparse_ca
tegorical_accuracy: 0.2543 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 1.8845 - sparse_categorical_accuracy: 0.2578 - val_loss: 1.9000 - val_sparse_ca
tegorical_accuracy: 0.2477 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 1.8724 - sparse_categorical_accuracy: 0.2610 - val_loss: 1.9016 - val_sparse_ca
tegorical_accuracy: 0.2440 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 1.8685 - sparse_categorical_accuracy: 0.2625 - val_loss: 1.8818 - val_sparse_ca
tegorical_accuracy: 0.2571 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 1.8710 - sparse_categorical_accuracy: 0.2633 - val_loss: 1.8945 - val_sparse_ca
tegorical_accuracy: 0.2748 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 1.8627 - sparse_categorical_accuracy: 0.2643 - val_loss: 1.8812 - val_sparse_ca
tegorical_accuracy: 0.2636 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 1.8694 - sparse_categorical_accuracy: 0.2645 - val_loss: 1.8874 - val_sparse_ca
tegorical_accuracy: 0.2727 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 1.8660 - sparse_categorical_accuracy: 0.2652 - val_loss: 1.8774 - val_sparse_ca
tegorical_accuracy: 0.2706 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 1.8595 - sparse_categorical_accuracy: 0.2666 - val_loss: 1.9135 - val_sparse_ca
tegorical_accuracy: 0.2643 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 1.8585 - sparse_categorical_accuracy: 0.2695 - val_loss: 1.8756 - val_sparse_ca
tegorical_accuracy: 0.2671 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 1.8636 - sparse_categorical_accuracy: 0.2677 - val_loss: 1.9071 - val_sparse_ca
tegorical_accuracy: 0.2566 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 1.8633 - sparse_categorical_accuracy: 0.2661 - val_loss: 1.8951 - val_sparse_ca
tegorical_accuracy: 0.2692 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 1.8630 - sparse_categorical_accuracy: 0.2675 - val_loss: 1.8804 - val_sparse_ca
tegorical_accuracy: 0.2770 - 1s/epoch - 2ms/step
Epoch 32/100
```

```
469/469 - 1s - loss: 1.8605 - sparse_categorical_accuracy: 0.2678 - val_loss: 1.9526 - val_sparse_ca
tegorical_accuracy: 0.2506 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 1.8705 - sparse_categorical_accuracy: 0.2654 - val_loss: 1.8675 - val_sparse_ca
tegorical_accuracy: 0.2836 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 1.8568 - sparse_categorical_accuracy: 0.2692 - val_loss: 1.9172 - val_sparse_ca
tegorical_accuracy: 0.2444 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 1.8503 - sparse_categorical_accuracy: 0.2695 - val_loss: 1.9614 - val_sparse_ca
tegorical_accuracy: 0.2322 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 1.8709 - sparse_categorical_accuracy: 0.2672 - val_loss: 1.8788 - val_sparse_ca
tegorical_accuracy: 0.2731 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 1.8559 - sparse_categorical_accuracy: 0.2689 - val_loss: 1.8671 - val_sparse_ca
tegorical_accuracy: 0.2847 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 1.8569 - sparse_categorical_accuracy: 0.2678 - val_loss: 1.8709 - val_sparse_ca
tegorical_accuracy: 0.2792 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 1.8576 - sparse_categorical_accuracy: 0.2694 - val_loss: 1.8704 - val_sparse_ca
tegorical_accuracy: 0.2775 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 1.8562 - sparse_categorical_accuracy: 0.2709 - val_loss: 1.8929 - val_sparse_ca
tegorical_accuracy: 0.2699 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 1.8591 - sparse_categorical_accuracy: 0.2692 - val_loss: 1.8722 - val_sparse_ca
tegorical_accuracy: 0.2883 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 1.8584 - sparse_categorical_accuracy: 0.2706 - val_loss: 1.8999 - val_sparse_ca
tegorical_accuracy: 0.2726 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 1.8571 - sparse_categorical_accuracy: 0.2707 - val_loss: 1.8864 - val_sparse_ca
tegorical_accuracy: 0.2800 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 1.8640 - sparse_categorical_accuracy: 0.2692 - val_loss: 1.8699 - val_sparse_ca
tegorical_accuracy: 0.2875 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 1.8605 - sparse_categorical_accuracy: 0.2707 - val_loss: 1.9003 - val_sparse_ca
tegorical_accuracy: 0.2725 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 1.8487 - sparse_categorical_accuracy: 0.2722 - val_loss: 1.8715 - val_sparse_ca
tegorical_accuracy: 0.2748 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 1.8566 - sparse_categorical_accuracy: 0.2713 - val_loss: 1.8871 - val_sparse_ca
tegorical_accuracy: 0.2591 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 1.8622 - sparse_categorical_accuracy: 0.2697 - val_loss: 1.8643 - val_sparse_ca
tegorical_accuracy: 0.2825 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 1.8563 - sparse_categorical_accuracy: 0.2698 - val_loss: 1.8606 - val_sparse_ca
tegorical_accuracy: 0.2720 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 1.8523 - sparse_categorical_accuracy: 0.2728 - val_loss: 1.8604 - val_sparse_ca
tegorical_accuracy: 0.2752 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 1.8616 - sparse_categorical_accuracy: 0.2712 - val_loss: 1.8563 - val_sparse_ca
tegorical_accuracy: 0.2878 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 1.8564 - sparse_categorical_accuracy: 0.2727 - val_loss: 1.8605 - val_sparse_ca
tegorical_accuracy: 0.2877 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 1.8582 - sparse_categorical_accuracy: 0.2699 - val_loss: 1.8754 - val_sparse_ca
tegorical_accuracy: 0.2757 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 1.8540 - sparse_categorical_accuracy: 0.2754 - val_loss: 1.8914 - val_sparse_ca
tegorical_accuracy: 0.2772 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 1.8516 - sparse_categorical_accuracy: 0.2726 - val_loss: 1.8614 - val_sparse_ca
tegorical_accuracy: 0.2830 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 1.8610 - sparse_categorical_accuracy: 0.2702 - val_loss: 1.9327 - val_sparse_ca
tegorical_accuracy: 0.2484 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 1.8579 - sparse_categorical_accuracy: 0.2717 - val_loss: 1.8828 - val_sparse_ca
tegorical_accuracy: 0.2739 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 1.8600 - sparse_categorical_accuracy: 0.2721 - val_loss: 1.8579 - val_sparse_ca
tegorical_accuracy: 0.2927 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 1.8503 - sparse_categorical_accuracy: 0.2743 - val_loss: 1.9259 - val_sparse_ca
tegorical_accuracy: 0.2558 - 1s/epoch - 2ms/step
```

```
Epoch 60/100
469/469 - 1s - loss: 1.8553 - sparse_categorical_accuracy: 0.2735 - val_loss: 1.8688 - val_sparse_ca
tegorical_accuracy: 0.2773 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 1.8526 - sparse_categorical_accuracy: 0.2722 - val_loss: 1.8625 - val_sparse_ca
tegorical_accuracy: 0.2863 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 1.8500 - sparse_categorical_accuracy: 0.2745 - val_loss: 1.8644 - val_sparse_ca
tegorical_accuracy: 0.2802 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 1.8566 - sparse_categorical_accuracy: 0.2710 - val_loss: 1.8665 - val_sparse_ca
tegorical_accuracy: 0.2860 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 1.8615 - sparse_categorical_accuracy: 0.2710 - val_loss: 1.8834 - val_sparse_ca
tegorical_accuracy: 0.2690 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 1.8571 - sparse_categorical_accuracy: 0.2709 - val_loss: 1.8992 - val_sparse_ca
tegorical_accuracy: 0.2622 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 1.8492 - sparse_categorical_accuracy: 0.2720 - val_loss: 1.9037 - val_sparse_ca
tegorical_accuracy: 0.2827 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 1.8547 - sparse_categorical_accuracy: 0.2725 - val_loss: 1.8736 - val_sparse_ca
tegorical_accuracy: 0.2752 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 1.8510 - sparse_categorical_accuracy: 0.2715 - val_loss: 1.8664 - val_sparse_ca
tegorical_accuracy: 0.2683 - 1s/epoch - 3ms/step
Epoch 69/100
469/469 - 1s - loss: 1.8531 - sparse_categorical_accuracy: 0.2726 - val_loss: 1.8616 - val_sparse_ca
tegorical_accuracy: 0.2853 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 1.8600 - sparse_categorical_accuracy: 0.2724 - val_loss: 1.8754 - val_sparse_ca
tegorical_accuracy: 0.2763 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 1.8558 - sparse_categorical_accuracy: 0.2720 - val_loss: 1.8710 - val_sparse_ca
tegorical_accuracy: 0.2874 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 1.8594 - sparse_categorical_accuracy: 0.2733 - val_loss: 1.8654 - val_sparse_ca
tegorical_accuracy: 0.2912 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 1.8465 - sparse_categorical_accuracy: 0.2730 - val_loss: 1.8567 - val_sparse_ca
tegorical_accuracy: 0.2796 - 1s/epoch - 3ms/step
Epoch 74/100
469/469 - 1s - loss: 1.8518 - sparse_categorical_accuracy: 0.2716 - val_loss: 1.8685 - val_sparse_ca
tegorical_accuracy: 0.2869 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 1.8559 - sparse_categorical_accuracy: 0.2731 - val_loss: 1.9197 - val_sparse_ca
tegorical_accuracy: 0.2652 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 1.8604 - sparse_categorical_accuracy: 0.2726 - val_loss: 1.8845 - val_sparse_ca
tegorical_accuracy: 0.2645 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 1.8480 - sparse_categorical_accuracy: 0.2740 - val_loss: 1.8650 - val_sparse_ca
tegorical_accuracy: 0.2873 - 1s/epoch - 3ms/step
Epoch 78/100
469/469 - 1s - loss: 1.8560 - sparse_categorical_accuracy: 0.2737 - val_loss: 1.8580 - val_sparse_ca
tegorical_accuracy: 0.2840 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 1.8544 - sparse_categorical_accuracy: 0.2716 - val_loss: 1.8753 - val_sparse_ca
tegorical_accuracy: 0.2861 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 1.8619 - sparse_categorical_accuracy: 0.2722 - val_loss: 1.9257 - val_sparse_ca
tegorical_accuracy: 0.2539 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 1.8532 - sparse_categorical_accuracy: 0.2722 - val_loss: 1.8563 - val_sparse_ca
tegorical_accuracy: 0.2877 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 1.8491 - sparse_categorical_accuracy: 0.2757 - val_loss: 1.8623 - val_sparse_ca
tegorical_accuracy: 0.2901 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 1.8548 - sparse_categorical_accuracy: 0.2726 - val_loss: 1.8683 - val_sparse_ca
tegorical_accuracy: 0.2771 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 1.8488 - sparse_categorical_accuracy: 0.2742 - val_loss: 1.8620 - val_sparse_ca
tegorical_accuracy: 0.2814 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 1.8509 - sparse_categorical_accuracy: 0.2752 - val_loss: 1.8579 - val_sparse_ca
tegorical_accuracy: 0.2811 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 1.8515 - sparse_categorical_accuracy: 0.2722 - val_loss: 1.8688 - val_sparse_ca
tegorical_accuracy: 0.2819 - 1s/epoch - 3ms/step
Epoch 87/100
469/469 - 1s - loss: 1.8510 - sparse_categorical_accuracy: 0.2749 - val_loss: 1.8533 - val_sparse_ca
```

```
tegorical_accuracy: 0.2841 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 1.8483 - sparse_categorical_accuracy: 0.2742 - val_loss: 1.8599 - val_sparse_ca
tegorical_accuracy: 0.2747 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 1.8499 - sparse_categorical_accuracy: 0.2745 - val_loss: 1.8664 - val_sparse_ca
tegorical_accuracy: 0.2845 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 1.8473 - sparse_categorical_accuracy: 0.2714 - val_loss: 1.8872 - val_sparse_ca
tegorical_accuracy: 0.2731 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 1.8483 - sparse_categorical_accuracy: 0.2745 - val_loss: 1.8654 - val_sparse_ca
tegorical_accuracy: 0.2715 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 1.8595 - sparse_categorical_accuracy: 0.2694 - val_loss: 1.8637 - val_sparse_ca
tegorical_accuracy: 0.2883 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 1.8422 - sparse_categorical_accuracy: 0.2745 - val_loss: 1.9703 - val_sparse_ca
tegorical_accuracy: 0.2373 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 1.8474 - sparse_categorical_accuracy: 0.2738 - val_loss: 1.9491 - val_sparse_ca
tegorical_accuracy: 0.2489 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 1.8491 - sparse_categorical_accuracy: 0.2734 - val_loss: 1.8632 - val_sparse_ca
tegorical_accuracy: 0.2762 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 1.8460 - sparse_categorical_accuracy: 0.2724 - val_loss: 1.8792 - val_sparse_ca
tegorical_accuracy: 0.2709 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 1.8558 - sparse_categorical_accuracy: 0.2735 - val_loss: 1.8751 - val_sparse_ca
tegorical_accuracy: 0.2784 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 1.8522 - sparse_categorical_accuracy: 0.2742 - val_loss: 1.8623 - val_sparse_ca
tegorical_accuracy: 0.2804 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 1.8524 - sparse_categorical_accuracy: 0.2737 - val_loss: 1.8810 - val_sparse_ca
tegorical_accuracy: 0.2833 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 1.8472 - sparse_categorical_accuracy: 0.2740 - val_loss: 1.8617 - val_sparse_ca
tegorical_accuracy: 0.2804 - 1s/epoch - 2ms/step
```

In [33]:

```
model2.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 1.8617 - sparse_categorical_accuracy: 0.2804 - 414ms/epoch - 1ms/step
```

Out[33]:

```
[1.861693024635315, 0.28040000796318054]
```

Model3: Regularization penalty is L1 with parameter 10^{-3}

In [34]:

```
model3 = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform',
  activity_regularizer = tf.keras.regularizers.L1(0.001)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
```

In [35]:

```
# Fix the learning rate to be 0.001
model3.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [36]:

```
history3 = model3.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 2s - loss: 1.3524 - sparse_categorical_accuracy: 0.5560 - val_loss: 1.0354 - val_sparse_ca
```

```
tegorical_accuracy: 0.6552 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 0.8993 - sparse_categorical_accuracy: 0.7054 - val_loss: 0.7858 - val_sparse_ca
tegorical_accuracy: 0.7628 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 0.6733 - sparse_categorical_accuracy: 0.8345 - val_loss: 0.6193 - val_sparse_ca
tegorical_accuracy: 0.8660 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 0.5782 - sparse_categorical_accuracy: 0.8765 - val_loss: 0.5841 - val_sparse_ca
tegorical_accuracy: 0.8728 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 0.5646 - sparse_categorical_accuracy: 0.8798 - val_loss: 0.5917 - val_sparse_ca
tegorical_accuracy: 0.8737 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 0.5507 - sparse_categorical_accuracy: 0.8839 - val_loss: 0.5809 - val_sparse_ca
tegorical_accuracy: 0.8831 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.5407 - sparse_categorical_accuracy: 0.8857 - val_loss: 0.5429 - val_sparse_ca
tegorical_accuracy: 0.8910 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 0.5445 - sparse_categorical_accuracy: 0.8857 - val_loss: 0.5467 - val_sparse_ca
tegorical_accuracy: 0.8827 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 0.5373 - sparse_categorical_accuracy: 0.8882 - val_loss: 0.5337 - val_sparse_ca
tegorical_accuracy: 0.8919 - 1s/epoch - 3ms/step
Epoch 10/100
469/469 - 1s - loss: 0.5329 - sparse_categorical_accuracy: 0.8898 - val_loss: 0.5349 - val_sparse_ca
tegorical_accuracy: 0.8886 - 1s/epoch - 3ms/step
Epoch 11/100
469/469 - 1s - loss: 0.5324 - sparse_categorical_accuracy: 0.8890 - val_loss: 0.5285 - val_sparse_ca
tegorical_accuracy: 0.8964 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.5334 - sparse_categorical_accuracy: 0.8894 - val_loss: 0.5344 - val_sparse_ca
tegorical_accuracy: 0.8940 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.5281 - sparse_categorical_accuracy: 0.8898 - val_loss: 0.5203 - val_sparse_ca
tegorical_accuracy: 0.8969 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.5254 - sparse_categorical_accuracy: 0.8909 - val_loss: 0.5714 - val_sparse_ca
tegorical_accuracy: 0.8784 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 0.5302 - sparse_categorical_accuracy: 0.8908 - val_loss: 0.5502 - val_sparse_ca
tegorical_accuracy: 0.8869 - 1s/epoch - 3ms/step
Epoch 16/100
469/469 - 1s - loss: 0.5329 - sparse_categorical_accuracy: 0.8910 - val_loss: 0.5368 - val_sparse_ca
tegorical_accuracy: 0.8936 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.5322 - sparse_categorical_accuracy: 0.8911 - val_loss: 0.5769 - val_sparse_ca
tegorical_accuracy: 0.8811 - 1s/epoch - 3ms/step
Epoch 18/100
469/469 - 1s - loss: 0.5444 - sparse_categorical_accuracy: 0.8873 - val_loss: 0.5332 - val_sparse_ca
tegorical_accuracy: 0.8946 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.5258 - sparse_categorical_accuracy: 0.8924 - val_loss: 0.6054 - val_sparse_ca
tegorical_accuracy: 0.8813 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.5481 - sparse_categorical_accuracy: 0.8895 - val_loss: 0.6253 - val_sparse_ca
tegorical_accuracy: 0.8795 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.5395 - sparse_categorical_accuracy: 0.8914 - val_loss: 0.5571 - val_sparse_ca
tegorical_accuracy: 0.8886 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.5407 - sparse_categorical_accuracy: 0.8900 - val_loss: 0.5337 - val_sparse_ca
tegorical_accuracy: 0.8971 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.5510 - sparse_categorical_accuracy: 0.8878 - val_loss: 0.5527 - val_sparse_ca
tegorical_accuracy: 0.8902 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.5512 - sparse_categorical_accuracy: 0.8874 - val_loss: 0.5402 - val_sparse_ca
tegorical_accuracy: 0.8949 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.5443 - sparse_categorical_accuracy: 0.8883 - val_loss: 0.5892 - val_sparse_ca
tegorical_accuracy: 0.8822 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.5452 - sparse_categorical_accuracy: 0.8892 - val_loss: 0.5498 - val_sparse_ca
tegorical_accuracy: 0.8915 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.5568 - sparse_categorical_accuracy: 0.8856 - val_loss: 0.5872 - val_sparse_ca
tegorical_accuracy: 0.8852 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.5427 - sparse_categorical_accuracy: 0.8900 - val_loss: 0.6178 - val_sparse_ca
tegorical_accuracy: 0.8803 - 1s/epoch - 2ms/step
Epoch 29/100
```

```
469/469 - 1s - loss: 0.5425 - sparse_categorical_accuracy: 0.8907 - val_loss: 0.5707 - val_sparse_ca
tegorical_accuracy: 0.8834 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.5546 - sparse_categorical_accuracy: 0.8859 - val_loss: 0.6032 - val_sparse_ca
tegorical_accuracy: 0.8740 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.5691 - sparse_categorical_accuracy: 0.8865 - val_loss: 0.5603 - val_sparse_ca
tegorical_accuracy: 0.8880 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.5652 - sparse_categorical_accuracy: 0.8867 - val_loss: 0.6513 - val_sparse_ca
tegorical_accuracy: 0.8659 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.5762 - sparse_categorical_accuracy: 0.8823 - val_loss: 0.6174 - val_sparse_ca
tegorical_accuracy: 0.8770 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 2s - loss: 0.5705 - sparse_categorical_accuracy: 0.8844 - val_loss: 0.6336 - val_sparse_ca
tegorical_accuracy: 0.8796 - 2s/epoch - 4ms/step
Epoch 35/100
469/469 - 1s - loss: 0.5847 - sparse_categorical_accuracy: 0.8827 - val_loss: 0.6057 - val_sparse_ca
tegorical_accuracy: 0.8807 - 1s/epoch - 3ms/step
Epoch 36/100
469/469 - 1s - loss: 0.5775 - sparse_categorical_accuracy: 0.8849 - val_loss: 0.6315 - val_sparse_ca
tegorical_accuracy: 0.8793 - 1s/epoch - 3ms/step
Epoch 37/100
469/469 - 2s - loss: 0.5679 - sparse_categorical_accuracy: 0.8864 - val_loss: 0.5552 - val_sparse_ca
tegorical_accuracy: 0.8938 - 2s/epoch - 4ms/step
Epoch 38/100
469/469 - 1s - loss: 0.5698 - sparse_categorical_accuracy: 0.8862 - val_loss: 0.6446 - val_sparse_ca
tegorical_accuracy: 0.8730 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 2s - loss: 0.5787 - sparse_categorical_accuracy: 0.8859 - val_loss: 0.5874 - val_sparse_ca
tegorical_accuracy: 0.8852 - 2s/epoch - 3ms/step
Epoch 40/100
469/469 - 1s - loss: 0.5772 - sparse_categorical_accuracy: 0.8872 - val_loss: 0.5807 - val_sparse_ca
tegorical_accuracy: 0.8877 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.5811 - sparse_categorical_accuracy: 0.8840 - val_loss: 0.5856 - val_sparse_ca
tegorical_accuracy: 0.8899 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.5961 - sparse_categorical_accuracy: 0.8833 - val_loss: 0.6076 - val_sparse_ca
tegorical_accuracy: 0.8844 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.5859 - sparse_categorical_accuracy: 0.8849 - val_loss: 0.6251 - val_sparse_ca
tegorical_accuracy: 0.8818 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.5960 - sparse_categorical_accuracy: 0.8829 - val_loss: 0.6522 - val_sparse_ca
tegorical_accuracy: 0.8680 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.6103 - sparse_categorical_accuracy: 0.8822 - val_loss: 0.6557 - val_sparse_ca
tegorical_accuracy: 0.8808 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.6132 - sparse_categorical_accuracy: 0.8845 - val_loss: 0.6721 - val_sparse_ca
tegorical_accuracy: 0.8770 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.6230 - sparse_categorical_accuracy: 0.8804 - val_loss: 0.7470 - val_sparse_ca
tegorical_accuracy: 0.8585 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.6106 - sparse_categorical_accuracy: 0.8841 - val_loss: 0.6350 - val_sparse_ca
tegorical_accuracy: 0.8742 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.6355 - sparse_categorical_accuracy: 0.8792 - val_loss: 0.6607 - val_sparse_ca
tegorical_accuracy: 0.8814 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.5983 - sparse_categorical_accuracy: 0.8867 - val_loss: 0.6010 - val_sparse_ca
tegorical_accuracy: 0.8857 - 1s/epoch - 3ms/step
Epoch 51/100
469/469 - 1s - loss: 0.5968 - sparse_categorical_accuracy: 0.8856 - val_loss: 0.6267 - val_sparse_ca
tegorical_accuracy: 0.8808 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.6094 - sparse_categorical_accuracy: 0.8816 - val_loss: 0.6517 - val_sparse_ca
tegorical_accuracy: 0.8730 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.6004 - sparse_categorical_accuracy: 0.8853 - val_loss: 0.6643 - val_sparse_ca
tegorical_accuracy: 0.8835 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.6007 - sparse_categorical_accuracy: 0.8837 - val_loss: 0.6163 - val_sparse_ca
tegorical_accuracy: 0.8833 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.6011 - sparse_categorical_accuracy: 0.8797 - val_loss: 0.6399 - val_sparse_ca
tegorical_accuracy: 0.8734 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.6124 - sparse_categorical_accuracy: 0.8795 - val_loss: 0.6189 - val_sparse_ca
tegorical_accuracy: 0.8784 - 1s/epoch - 2ms/step
```

```
Epoch 57/100
469/469 - 1s - loss: 0.6203 - sparse_categorical_accuracy: 0.8778 - val_loss: 0.6251 - val_sparse_ca
tegorical_accuracy: 0.8836 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.6228 - sparse_categorical_accuracy: 0.8781 - val_loss: 0.6534 - val_sparse_ca
tegorical_accuracy: 0.8788 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.6173 - sparse_categorical_accuracy: 0.8806 - val_loss: 0.6272 - val_sparse_ca
tegorical_accuracy: 0.8716 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.6242 - sparse_categorical_accuracy: 0.8787 - val_loss: 0.6417 - val_sparse_ca
tegorical_accuracy: 0.8839 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.6323 - sparse_categorical_accuracy: 0.8786 - val_loss: 0.6325 - val_sparse_ca
tegorical_accuracy: 0.8812 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.6235 - sparse_categorical_accuracy: 0.8784 - val_loss: 0.6092 - val_sparse_ca
tegorical_accuracy: 0.8836 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 0.5957 - sparse_categorical_accuracy: 0.8834 - val_loss: 0.6655 - val_sparse_ca
tegorical_accuracy: 0.8718 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.6107 - sparse_categorical_accuracy: 0.8825 - val_loss: 0.6299 - val_sparse_ca
tegorical_accuracy: 0.8782 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.6537 - sparse_categorical_accuracy: 0.8686 - val_loss: 0.6514 - val_sparse_ca
tegorical_accuracy: 0.8739 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.6548 - sparse_categorical_accuracy: 0.8755 - val_loss: 0.6879 - val_sparse_ca
tegorical_accuracy: 0.8782 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.6356 - sparse_categorical_accuracy: 0.8798 - val_loss: 0.6584 - val_sparse_ca
tegorical_accuracy: 0.8797 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 0.6478 - sparse_categorical_accuracy: 0.8738 - val_loss: 0.6392 - val_sparse_ca
tegorical_accuracy: 0.8813 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 0.6319 - sparse_categorical_accuracy: 0.8804 - val_loss: 0.6647 - val_sparse_ca
tegorical_accuracy: 0.8766 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.6214 - sparse_categorical_accuracy: 0.8797 - val_loss: 0.7469 - val_sparse_ca
tegorical_accuracy: 0.8606 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.6298 - sparse_categorical_accuracy: 0.8814 - val_loss: 0.6814 - val_sparse_ca
tegorical_accuracy: 0.8725 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.6742 - sparse_categorical_accuracy: 0.8726 - val_loss: 0.6878 - val_sparse_ca
tegorical_accuracy: 0.8734 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.6577 - sparse_categorical_accuracy: 0.8698 - val_loss: 0.6493 - val_sparse_ca
tegorical_accuracy: 0.8712 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.6362 - sparse_categorical_accuracy: 0.8767 - val_loss: 0.6591 - val_sparse_ca
tegorical_accuracy: 0.8750 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.6528 - sparse_categorical_accuracy: 0.8750 - val_loss: 0.6693 - val_sparse_ca
tegorical_accuracy: 0.8710 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.6822 - sparse_categorical_accuracy: 0.8715 - val_loss: 0.6896 - val_sparse_ca
tegorical_accuracy: 0.8833 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.6487 - sparse_categorical_accuracy: 0.8798 - val_loss: 0.6787 - val_sparse_ca
tegorical_accuracy: 0.8730 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.6570 - sparse_categorical_accuracy: 0.8795 - val_loss: 0.6579 - val_sparse_ca
tegorical_accuracy: 0.8839 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.6289 - sparse_categorical_accuracy: 0.8823 - val_loss: 0.6594 - val_sparse_ca
tegorical_accuracy: 0.8836 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.6461 - sparse_categorical_accuracy: 0.8749 - val_loss: 0.7325 - val_sparse_ca
tegorical_accuracy: 0.8665 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.6656 - sparse_categorical_accuracy: 0.8746 - val_loss: 0.6537 - val_sparse_ca
tegorical_accuracy: 0.8774 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.6517 - sparse_categorical_accuracy: 0.8767 - val_loss: 0.7053 - val_sparse_ca
tegorical_accuracy: 0.8590 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.6728 - sparse_categorical_accuracy: 0.8736 - val_loss: 0.6756 - val_sparse_ca
tegorical_accuracy: 0.8660 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.6454 - sparse_categorical_accuracy: 0.8779 - val_loss: 0.6981 - val_sparse_ca
```

```
tegorical_accuracy: 0.8648 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.6723 - sparse_categorical_accuracy: 0.8713 - val_loss: 0.7286 - val_sparse_ca
tegorical_accuracy: 0.8691 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.6881 - sparse_categorical_accuracy: 0.8726 - val_loss: 0.6926 - val_sparse_ca
tegorical_accuracy: 0.8799 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.6881 - sparse_categorical_accuracy: 0.8718 - val_loss: 0.7024 - val_sparse_ca
tegorical_accuracy: 0.8831 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.6482 - sparse_categorical_accuracy: 0.8824 - val_loss: 0.6764 - val_sparse_ca
tegorical_accuracy: 0.8760 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.6384 - sparse_categorical_accuracy: 0.8821 - val_loss: 0.6664 - val_sparse_ca
tegorical_accuracy: 0.8802 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 2s - loss: 0.6402 - sparse_categorical_accuracy: 0.8777 - val_loss: 0.6943 - val_sparse_ca
tegorical_accuracy: 0.8784 - 2s/epoch - 3ms/step
Epoch 91/100
469/469 - 2s - loss: 0.6727 - sparse_categorical_accuracy: 0.8684 - val_loss: 0.6711 - val_sparse_ca
tegorical_accuracy: 0.8777 - 2s/epoch - 3ms/step
Epoch 92/100
469/469 - 1s - loss: 0.6626 - sparse_categorical_accuracy: 0.8727 - val_loss: 0.6875 - val_sparse_ca
tegorical_accuracy: 0.8657 - 1s/epoch - 3ms/step
Epoch 93/100
469/469 - 2s - loss: 0.7006 - sparse_categorical_accuracy: 0.8748 - val_loss: 0.7136 - val_sparse_ca
tegorical_accuracy: 0.8705 - 2s/epoch - 4ms/step
Epoch 94/100
469/469 - 1s - loss: 0.6686 - sparse_categorical_accuracy: 0.8780 - val_loss: 0.6658 - val_sparse_ca
tegorical_accuracy: 0.8817 - 1s/epoch - 3ms/step
Epoch 95/100
469/469 - 2s - loss: 0.6669 - sparse_categorical_accuracy: 0.8787 - val_loss: 0.6822 - val_sparse_ca
tegorical_accuracy: 0.8803 - 2s/epoch - 3ms/step
Epoch 96/100
469/469 - 1s - loss: 0.6664 - sparse_categorical_accuracy: 0.8755 - val_loss: 0.6705 - val_sparse_ca
tegorical_accuracy: 0.8774 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.6817 - sparse_categorical_accuracy: 0.8749 - val_loss: 0.6782 - val_sparse_ca
tegorical_accuracy: 0.8782 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.6567 - sparse_categorical_accuracy: 0.8766 - val_loss: 0.6914 - val_sparse_ca
tegorical_accuracy: 0.8718 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.6623 - sparse_categorical_accuracy: 0.8753 - val_loss: 0.7127 - val_sparse_ca
tegorical_accuracy: 0.8672 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.6538 - sparse_categorical_accuracy: 0.8783 - val_loss: 0.6666 - val_sparse_ca
tegorical_accuracy: 0.8689 - 1s/epoch - 2ms/step
```

In [37]:

```
model3.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.6666 - sparse_categorical_accuracy: 0.8689 - 394ms/epoch - 1ms/step
```

Out[37]:

```
[0.6665586233139038, 0.8689000010490417]
```

Model4: Regularization penalty is L2 with parameter 10^{-2}

In [38]:

```
model4 = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform',
  activity_regularizer = tf.keras.regularizers.L2(0.01)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
```

In [39]:

```
# Fix the learning rate to be 0.001
model4.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [40]:

```
history4 = model4.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 2s - loss: 3.2833 - sparse_categorical_accuracy: 0.1123 - val_loss: 2.3015 - val_sparse_ca
tegorical_accuracy: 0.1135 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 2.3018 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3014 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 2.3015 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3013 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 2.3014 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3013 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 2.3014 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3012 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 3ms/step
Epoch 14/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 2s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 2s/epoch - 4ms/step
Epoch 22/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
```

```
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 53/100
```

```
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
```

```
Epoch 81/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 2.3012 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3010 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 2.3013 - sparse_categorical_accuracy: 0.1124 - val_loss: 2.3011 - val_sparse_ca
tegorical_accuracy: 0.1135 - 1s/epoch - 2ms/step
```

In [41]:

```
model4.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 2.3011 - sparse_categorical_accuracy: 0.1135 - 405ms/epoch - 1ms/step
```

Out[41]:

```
[2.301110029220581, 0.11349999904632568]
```

Model5: Regularization penalty is L2 with parameter 10^{-4}

In [45]:

```
model5 = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform',
  activity_regularizer = tf.keras.regularizers.L2(0.0001)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
```

In [46]:

```
# Fix the learning rate to be 0.001
model5.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [47]:

```
history5 = model5.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 2s - loss: 2.1907 - sparse_categorical_accuracy: 0.1667 - val_loss: 2.0387 - val_sparse_ca
tegorical_accuracy: 0.2119 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 2.0278 - sparse_categorical_accuracy: 0.2154 - val_loss: 1.9299 - val_sparse_ca
tegorical_accuracy: 0.2646 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 1.8369 - sparse_categorical_accuracy: 0.3020 - val_loss: 1.8102 - val_sparse_ca
tegorical_accuracy: 0.2984 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 1.6989 - sparse_categorical_accuracy: 0.3609 - val_loss: 1.6227 - val_sparse_ca
tegorical_accuracy: 0.3615 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 1.5751 - sparse_categorical_accuracy: 0.3889 - val_loss: 1.5993 - val_sparse_ca
tegorical_accuracy: 0.3797 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 1.5331 - sparse_categorical_accuracy: 0.4039 - val_loss: 1.6061 - val_sparse_ca
tegorical_accuracy: 0.3787 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 1.4961 - sparse_categorical_accuracy: 0.4241 - val_loss: 1.5087 - val_sparse_ca
tegorical_accuracy: 0.4308 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 1.4826 - sparse_categorical_accuracy: 0.4368 - val_loss: 1.4765 - val_sparse_ca
tegorical_accuracy: 0.4364 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 1.4730 - sparse_categorical_accuracy: 0.4495 - val_loss: 1.5124 - val_sparse_ca
tegorical_accuracy: 0.4234 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 1.3985 - sparse_categorical_accuracy: 0.5058 - val_loss: 1.3115 - val_sparse_ca
tegorical_accuracy: 0.5656 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 1.3061 - sparse_categorical_accuracy: 0.5484 - val_loss: 1.2759 - val_sparse_ca
tegorical_accuracy: 0.5482 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 1.2812 - sparse_categorical_accuracy: 0.5607 - val_loss: 1.3043 - val_sparse_ca
tegorical_accuracy: 0.5629 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 1.2799 - sparse_categorical_accuracy: 0.5613 - val_loss: 1.2819 - val_sparse_ca
tegorical_accuracy: 0.5580 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 1.2912 - sparse_categorical_accuracy: 0.5601 - val_loss: 1.3713 - val_sparse_ca
tegorical_accuracy: 0.5650 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 1.3032 - sparse_categorical_accuracy: 0.5597 - val_loss: 1.3773 - val_sparse_ca
tegorical_accuracy: 0.5393 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 1.3066 - sparse_categorical_accuracy: 0.5584 - val_loss: 1.3490 - val_sparse_ca
tegorical_accuracy: 0.5531 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 1.3252 - sparse_categorical_accuracy: 0.5513 - val_loss: 1.3369 - val_sparse_ca
tegorical_accuracy: 0.5580 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 1.3273 - sparse_categorical_accuracy: 0.5561 - val_loss: 1.4450 - val_sparse_ca
tegorical_accuracy: 0.5317 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 1.3280 - sparse_categorical_accuracy: 0.5546 - val_loss: 1.3553 - val_sparse_ca
tegorical_accuracy: 0.5593 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 1.3539 - sparse_categorical_accuracy: 0.5486 - val_loss: 1.2960 - val_sparse_ca
tegorical_accuracy: 0.5606 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 1.3603 - sparse_categorical_accuracy: 0.5477 - val_loss: 1.3526 - val_sparse_ca
tegorical_accuracy: 0.5439 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 1.3601 - sparse_categorical_accuracy: 0.5448 - val_loss: 1.3247 - val_sparse_ca
tegorical_accuracy: 0.5587 - 1s/epoch - 3ms/step
```

```
Epoch 23/100
469/469 - 1s - loss: 1.3865 - sparse_categorical_accuracy: 0.5370 - val_loss: 1.3591 - val_sparse_ca
tegorical_accuracy: 0.5414 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 1.3640 - sparse_categorical_accuracy: 0.5362 - val_loss: 1.5587 - val_sparse_ca
tegorical_accuracy: 0.4980 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 1.4209 - sparse_categorical_accuracy: 0.5319 - val_loss: 1.3392 - val_sparse_ca
tegorical_accuracy: 0.5569 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 1.4184 - sparse_categorical_accuracy: 0.5321 - val_loss: 1.4783 - val_sparse_ca
tegorical_accuracy: 0.5218 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 1.4185 - sparse_categorical_accuracy: 0.5288 - val_loss: 1.3923 - val_sparse_ca
tegorical_accuracy: 0.5412 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 1.4560 - sparse_categorical_accuracy: 0.5148 - val_loss: 1.4388 - val_sparse_ca
tegorical_accuracy: 0.4941 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 1.4681 - sparse_categorical_accuracy: 0.5006 - val_loss: 1.4627 - val_sparse_ca
tegorical_accuracy: 0.5354 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 1.4757 - sparse_categorical_accuracy: 0.5098 - val_loss: 1.4250 - val_sparse_ca
tegorical_accuracy: 0.4915 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 1.4853 - sparse_categorical_accuracy: 0.5027 - val_loss: 1.4194 - val_sparse_ca
tegorical_accuracy: 0.4989 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 1.5168 - sparse_categorical_accuracy: 0.4894 - val_loss: 1.4312 - val_sparse_ca
tegorical_accuracy: 0.5209 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 1.5301 - sparse_categorical_accuracy: 0.4937 - val_loss: 1.4864 - val_sparse_ca
tegorical_accuracy: 0.4889 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 1.5391 - sparse_categorical_accuracy: 0.4848 - val_loss: 1.5129 - val_sparse_ca
tegorical_accuracy: 0.4749 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 1.5584 - sparse_categorical_accuracy: 0.4879 - val_loss: 1.5985 - val_sparse_ca
tegorical_accuracy: 0.4582 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 1.5608 - sparse_categorical_accuracy: 0.4783 - val_loss: 1.4987 - val_sparse_ca
tegorical_accuracy: 0.4895 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 1.5592 - sparse_categorical_accuracy: 0.4624 - val_loss: 1.6530 - val_sparse_ca
tegorical_accuracy: 0.4790 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 1.5674 - sparse_categorical_accuracy: 0.4708 - val_loss: 1.5760 - val_sparse_ca
tegorical_accuracy: 0.4394 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 1.6173 - sparse_categorical_accuracy: 0.4753 - val_loss: 1.6035 - val_sparse_ca
tegorical_accuracy: 0.4428 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 1.5981 - sparse_categorical_accuracy: 0.4591 - val_loss: 1.5454 - val_sparse_ca
tegorical_accuracy: 0.4513 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 1.6900 - sparse_categorical_accuracy: 0.4627 - val_loss: 1.6279 - val_sparse_ca
tegorical_accuracy: 0.4173 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 1.6834 - sparse_categorical_accuracy: 0.4613 - val_loss: 1.6240 - val_sparse_ca
tegorical_accuracy: 0.4571 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 1.6768 - sparse_categorical_accuracy: 0.4616 - val_loss: 1.6486 - val_sparse_ca
tegorical_accuracy: 0.4664 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 1.6562 - sparse_categorical_accuracy: 0.4720 - val_loss: 1.5637 - val_sparse_ca
tegorical_accuracy: 0.4880 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 1.6237 - sparse_categorical_accuracy: 0.4635 - val_loss: 1.7156 - val_sparse_ca
tegorical_accuracy: 0.4940 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 1.6900 - sparse_categorical_accuracy: 0.4608 - val_loss: 1.8045 - val_sparse_ca
tegorical_accuracy: 0.4269 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 1.6961 - sparse_categorical_accuracy: 0.4490 - val_loss: 1.8078 - val_sparse_ca
tegorical_accuracy: 0.3723 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 1.7477 - sparse_categorical_accuracy: 0.4578 - val_loss: 1.7746 - val_sparse_ca
tegorical_accuracy: 0.4436 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 1.7154 - sparse_categorical_accuracy: 0.4301 - val_loss: 1.7135 - val_sparse_ca
tegorical_accuracy: 0.4181 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 1.7777 - sparse_categorical_accuracy: 0.4303 - val_loss: 1.7023 - val_sparse_ca
```

```
tegorical_accuracy: 0.4266 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 1.7275 - sparse_categorical_accuracy: 0.4138 - val_loss: 1.7249 - val_sparse_ca
tegorical_accuracy: 0.3817 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 1.7569 - sparse_categorical_accuracy: 0.4211 - val_loss: 1.8275 - val_sparse_ca
tegorical_accuracy: 0.4315 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 1.7409 - sparse_categorical_accuracy: 0.4118 - val_loss: 1.7510 - val_sparse_ca
tegorical_accuracy: 0.4006 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 1.7475 - sparse_categorical_accuracy: 0.4079 - val_loss: 1.7194 - val_sparse_ca
tegorical_accuracy: 0.4034 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 1.8126 - sparse_categorical_accuracy: 0.3835 - val_loss: 1.7497 - val_sparse_ca
tegorical_accuracy: 0.3666 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 1.8146 - sparse_categorical_accuracy: 0.3604 - val_loss: 1.8113 - val_sparse_ca
tegorical_accuracy: 0.3449 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 1.8211 - sparse_categorical_accuracy: 0.4015 - val_loss: 1.8030 - val_sparse_ca
tegorical_accuracy: 0.3668 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 1.8461 - sparse_categorical_accuracy: 0.3737 - val_loss: 1.7701 - val_sparse_ca
tegorical_accuracy: 0.3862 - 1s/epoch - 3ms/step
Epoch 59/100
469/469 - 1s - loss: 1.8456 - sparse_categorical_accuracy: 0.3667 - val_loss: 1.8459 - val_sparse_ca
tegorical_accuracy: 0.3107 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 1.8845 - sparse_categorical_accuracy: 0.3901 - val_loss: 1.8432 - val_sparse_ca
tegorical_accuracy: 0.3986 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 1.8156 - sparse_categorical_accuracy: 0.3795 - val_loss: 1.8112 - val_sparse_ca
tegorical_accuracy: 0.4350 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 1.8529 - sparse_categorical_accuracy: 0.3956 - val_loss: 1.7773 - val_sparse_ca
tegorical_accuracy: 0.3853 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 1.8340 - sparse_categorical_accuracy: 0.3715 - val_loss: 1.8293 - val_sparse_ca
tegorical_accuracy: 0.3317 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 1.8752 - sparse_categorical_accuracy: 0.3642 - val_loss: 1.8534 - val_sparse_ca
tegorical_accuracy: 0.3204 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 1.9438 - sparse_categorical_accuracy: 0.3696 - val_loss: 1.8307 - val_sparse_ca
tegorical_accuracy: 0.3308 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 1.8891 - sparse_categorical_accuracy: 0.3678 - val_loss: 1.9528 - val_sparse_ca
tegorical_accuracy: 0.3642 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 1.8859 - sparse_categorical_accuracy: 0.3569 - val_loss: 1.9429 - val_sparse_ca
tegorical_accuracy: 0.3396 - 1s/epoch - 3ms/step
Epoch 68/100
469/469 - 1s - loss: 1.9224 - sparse_categorical_accuracy: 0.3422 - val_loss: 1.8134 - val_sparse_ca
tegorical_accuracy: 0.3595 - 1s/epoch - 3ms/step
Epoch 69/100
469/469 - 2s - loss: 1.9193 - sparse_categorical_accuracy: 0.3618 - val_loss: 2.6564 - val_sparse_ca
tegorical_accuracy: 0.3667 - 2s/epoch - 4ms/step
Epoch 70/100
469/469 - 2s - loss: 1.8997 - sparse_categorical_accuracy: 0.3476 - val_loss: 1.9544 - val_sparse_ca
tegorical_accuracy: 0.2970 - 2s/epoch - 4ms/step
Epoch 71/100
469/469 - 1s - loss: 1.9776 - sparse_categorical_accuracy: 0.3119 - val_loss: 2.1919 - val_sparse_ca
tegorical_accuracy: 0.3779 - 1s/epoch - 3ms/step
Epoch 72/100
469/469 - 1s - loss: 1.9371 - sparse_categorical_accuracy: 0.3365 - val_loss: 1.9030 - val_sparse_ca
tegorical_accuracy: 0.3529 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 1.9812 - sparse_categorical_accuracy: 0.3210 - val_loss: 1.9313 - val_sparse_ca
tegorical_accuracy: 0.3065 - 1s/epoch - 3ms/step
Epoch 74/100
469/469 - 1s - loss: 1.9449 - sparse_categorical_accuracy: 0.3304 - val_loss: 1.9103 - val_sparse_ca
tegorical_accuracy: 0.3239 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 1.9327 - sparse_categorical_accuracy: 0.3323 - val_loss: 1.8988 - val_sparse_ca
tegorical_accuracy: 0.3502 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 1.9603 - sparse_categorical_accuracy: 0.3085 - val_loss: 1.9915 - val_sparse_ca
tegorical_accuracy: 0.2879 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 1.9517 - sparse_categorical_accuracy: 0.3123 - val_loss: 1.9036 - val_sparse_ca
tegorical_accuracy: 0.3094 - 1s/epoch - 2ms/step
Epoch 78/100
```

```
469/469 - 1s - loss: 2.0228 - sparse_categorical_accuracy: 0.3203 - val_loss: 2.0004 - val_sparse_ca
tegorical_accuracy: 0.3189 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 2.0407 - sparse_categorical_accuracy: 0.2983 - val_loss: 2.5721 - val_sparse_ca
tegorical_accuracy: 0.3945 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 2.0302 - sparse_categorical_accuracy: 0.3136 - val_loss: 2.0401 - val_sparse_ca
tegorical_accuracy: 0.3150 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 2.0187 - sparse_categorical_accuracy: 0.2935 - val_loss: 1.9803 - val_sparse_ca
tegorical_accuracy: 0.3145 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 2.0032 - sparse_categorical_accuracy: 0.2891 - val_loss: 1.9839 - val_sparse_ca
tegorical_accuracy: 0.2584 - 1s/epoch - 3ms/step
Epoch 83/100
469/469 - 1s - loss: 2.0371 - sparse_categorical_accuracy: 0.2571 - val_loss: 1.9855 - val_sparse_ca
tegorical_accuracy: 0.2421 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 2.0758 - sparse_categorical_accuracy: 0.2560 - val_loss: 2.1343 - val_sparse_ca
tegorical_accuracy: 0.2860 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 2.0366 - sparse_categorical_accuracy: 0.2510 - val_loss: 1.9988 - val_sparse_ca
tegorical_accuracy: 0.2565 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 2.0460 - sparse_categorical_accuracy: 0.2435 - val_loss: 2.1060 - val_sparse_ca
tegorical_accuracy: 0.2175 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 2.1235 - sparse_categorical_accuracy: 0.2630 - val_loss: 2.0925 - val_sparse_ca
tegorical_accuracy: 0.2399 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 2.0370 - sparse_categorical_accuracy: 0.2474 - val_loss: 1.9859 - val_sparse_ca
tegorical_accuracy: 0.3048 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 2.0711 - sparse_categorical_accuracy: 0.2903 - val_loss: 2.0002 - val_sparse_ca
tegorical_accuracy: 0.2788 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 2.0436 - sparse_categorical_accuracy: 0.2551 - val_loss: 2.0263 - val_sparse_ca
tegorical_accuracy: 0.2489 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 2.0461 - sparse_categorical_accuracy: 0.2283 - val_loss: 2.1188 - val_sparse_ca
tegorical_accuracy: 0.2395 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 2.1219 - sparse_categorical_accuracy: 0.2329 - val_loss: 2.0517 - val_sparse_ca
tegorical_accuracy: 0.2291 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 2.0820 - sparse_categorical_accuracy: 0.2183 - val_loss: 2.0362 - val_sparse_ca
tegorical_accuracy: 0.2380 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 2.1359 - sparse_categorical_accuracy: 0.2347 - val_loss: 2.1170 - val_sparse_ca
tegorical_accuracy: 0.2415 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 2.1283 - sparse_categorical_accuracy: 0.2192 - val_loss: 2.1455 - val_sparse_ca
tegorical_accuracy: 0.2182 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 2.1429 - sparse_categorical_accuracy: 0.2211 - val_loss: 2.1089 - val_sparse_ca
tegorical_accuracy: 0.1939 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 2.1280 - sparse_categorical_accuracy: 0.2059 - val_loss: 2.0768 - val_sparse_ca
tegorical_accuracy: 0.1978 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 2.1339 - sparse_categorical_accuracy: 0.1995 - val_loss: 2.0539 - val_sparse_ca
tegorical_accuracy: 0.2239 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 2.1272 - sparse_categorical_accuracy: 0.2097 - val_loss: 2.2624 - val_sparse_ca
tegorical_accuracy: 0.2235 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 2.1477 - sparse_categorical_accuracy: 0.2014 - val_loss: 2.1334 - val_sparse_ca
tegorical_accuracy: 0.1772 - 1s/epoch - 2ms/step
```

In [48]:

```
model5.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 2.1334 - sparse_categorical_accuracy: 0.1772 - 394ms/epoch - 1ms/step
```

Out[48]:

```
[2.13344144821167, 0.17720000445842743]
```

Model6: Regularization penalty is L2 with parameter 10^{-5}

In [50]:

```
model6 = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform',
  activity_regularizer = tf.keras.regularizers.L2(0.00001)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
```

In [51]:

```
model6.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [52]:

```
history6 = model6.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 2s - loss: 1.1305 - sparse_categorical_accuracy: 0.6603 - val_loss: 0.8640 - val_sparse_ca
tegorical_accuracy: 0.7444 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 0.6250 - sparse_categorical_accuracy: 0.8388 - val_loss: 0.5116 - val_sparse_ca
tegorical_accuracy: 0.8938 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 0.4721 - sparse_categorical_accuracy: 0.8987 - val_loss: 0.4684 - val_sparse_ca
tegorical_accuracy: 0.9007 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 0.4148 - sparse_categorical_accuracy: 0.9119 - val_loss: 0.4269 - val_sparse_ca
tegorical_accuracy: 0.9087 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 0.3934 - sparse_categorical_accuracy: 0.9183 - val_loss: 0.4308 - val_sparse_ca
tegorical_accuracy: 0.9146 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 0.3830 - sparse_categorical_accuracy: 0.9219 - val_loss: 0.4226 - val_sparse_ca
tegorical_accuracy: 0.9099 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.3727 - sparse_categorical_accuracy: 0.9225 - val_loss: 0.3986 - val_sparse_ca
tegorical_accuracy: 0.9215 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 0.3741 - sparse_categorical_accuracy: 0.9246 - val_loss: 0.4062 - val_sparse_ca
tegorical_accuracy: 0.9209 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 0.3667 - sparse_categorical_accuracy: 0.9252 - val_loss: 0.3775 - val_sparse_ca
tegorical_accuracy: 0.9220 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 0.3714 - sparse_categorical_accuracy: 0.9239 - val_loss: 0.4192 - val_sparse_ca
tegorical_accuracy: 0.9226 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 0.3623 - sparse_categorical_accuracy: 0.9264 - val_loss: 0.3910 - val_sparse_ca
tegorical_accuracy: 0.9166 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.3634 - sparse_categorical_accuracy: 0.9269 - val_loss: 0.4193 - val_sparse_ca
tegorical_accuracy: 0.9142 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.3667 - sparse_categorical_accuracy: 0.9262 - val_loss: 0.3893 - val_sparse_ca
tegorical_accuracy: 0.9198 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.3733 - sparse_categorical_accuracy: 0.9260 - val_loss: 0.3995 - val_sparse_ca
tegorical_accuracy: 0.9212 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 0.3625 - sparse_categorical_accuracy: 0.9262 - val_loss: 0.4097 - val_sparse_ca
tegorical_accuracy: 0.9126 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.3623 - sparse_categorical_accuracy: 0.9270 - val_loss: 0.5330 - val_sparse_ca
tegorical_accuracy: 0.8905 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.3680 - sparse_categorical_accuracy: 0.9264 - val_loss: 0.4900 - val_sparse_ca
tegorical_accuracy: 0.9042 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 0.3772 - sparse_categorical_accuracy: 0.9244 - val_loss: 0.4099 - val_sparse_ca
tegorical_accuracy: 0.9192 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.3881 - sparse_categorical_accuracy: 0.9212 - val_loss: 0.4135 - val_sparse_ca
```

```
tegorical_accuracy: 0.9165 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.3868 - sparse_categorical_accuracy: 0.9197 - val_loss: 0.4111 - val_sparse_ca
tegorical_accuracy: 0.9122 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.3945 - sparse_categorical_accuracy: 0.9226 - val_loss: 0.4153 - val_sparse_ca
tegorical_accuracy: 0.9173 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.3933 - sparse_categorical_accuracy: 0.9215 - val_loss: 0.4137 - val_sparse_ca
tegorical_accuracy: 0.9207 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.3935 - sparse_categorical_accuracy: 0.9228 - val_loss: 0.4141 - val_sparse_ca
tegorical_accuracy: 0.9190 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.3738 - sparse_categorical_accuracy: 0.9247 - val_loss: 0.4252 - val_sparse_ca
tegorical_accuracy: 0.9139 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.4067 - sparse_categorical_accuracy: 0.9201 - val_loss: 0.4995 - val_sparse_ca
tegorical_accuracy: 0.9113 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.4036 - sparse_categorical_accuracy: 0.9202 - val_loss: 0.4140 - val_sparse_ca
tegorical_accuracy: 0.9116 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.3988 - sparse_categorical_accuracy: 0.9205 - val_loss: 0.4524 - val_sparse_ca
tegorical_accuracy: 0.9050 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.4201 - sparse_categorical_accuracy: 0.9179 - val_loss: 0.4627 - val_sparse_ca
tegorical_accuracy: 0.9081 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.4233 - sparse_categorical_accuracy: 0.9184 - val_loss: 0.4687 - val_sparse_ca
tegorical_accuracy: 0.9075 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.4123 - sparse_categorical_accuracy: 0.9203 - val_loss: 0.5424 - val_sparse_ca
tegorical_accuracy: 0.9054 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.4513 - sparse_categorical_accuracy: 0.9135 - val_loss: 0.5054 - val_sparse_ca
tegorical_accuracy: 0.9065 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.4338 - sparse_categorical_accuracy: 0.9205 - val_loss: 0.4592 - val_sparse_ca
tegorical_accuracy: 0.9131 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.4498 - sparse_categorical_accuracy: 0.9165 - val_loss: 0.5027 - val_sparse_ca
tegorical_accuracy: 0.9015 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.4562 - sparse_categorical_accuracy: 0.9157 - val_loss: 0.4852 - val_sparse_ca
tegorical_accuracy: 0.9077 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.4448 - sparse_categorical_accuracy: 0.9171 - val_loss: 0.5269 - val_sparse_ca
tegorical_accuracy: 0.8914 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.4612 - sparse_categorical_accuracy: 0.9144 - val_loss: 0.4895 - val_sparse_ca
tegorical_accuracy: 0.9054 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.4550 - sparse_categorical_accuracy: 0.9148 - val_loss: 0.5179 - val_sparse_ca
tegorical_accuracy: 0.8976 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.4816 - sparse_categorical_accuracy: 0.9124 - val_loss: 0.4834 - val_sparse_ca
tegorical_accuracy: 0.9109 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.4840 - sparse_categorical_accuracy: 0.9114 - val_loss: 0.4774 - val_sparse_ca
tegorical_accuracy: 0.9124 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 0.4783 - sparse_categorical_accuracy: 0.9132 - val_loss: 0.5077 - val_sparse_ca
tegorical_accuracy: 0.9087 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.4964 - sparse_categorical_accuracy: 0.9115 - val_loss: 0.5287 - val_sparse_ca
tegorical_accuracy: 0.9018 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.4876 - sparse_categorical_accuracy: 0.9140 - val_loss: 0.5326 - val_sparse_ca
tegorical_accuracy: 0.9075 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.4892 - sparse_categorical_accuracy: 0.9102 - val_loss: 0.5005 - val_sparse_ca
tegorical_accuracy: 0.9009 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.4955 - sparse_categorical_accuracy: 0.9102 - val_loss: 0.5362 - val_sparse_ca
tegorical_accuracy: 0.8958 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.4879 - sparse_categorical_accuracy: 0.9112 - val_loss: 0.5337 - val_sparse_ca
tegorical_accuracy: 0.9064 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.4811 - sparse_categorical_accuracy: 0.9125 - val_loss: 0.5611 - val_sparse_ca
tegorical_accuracy: 0.8900 - 1s/epoch - 2ms/step
Epoch 47/100
```

```
469/469 - 1s - loss: 0.5169 - sparse_categorical_accuracy: 0.9066 - val_loss: 0.5381 - val_sparse_ca
tegorical_accuracy: 0.9096 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.5063 - sparse_categorical_accuracy: 0.9060 - val_loss: 0.5481 - val_sparse_ca
tegorical_accuracy: 0.8989 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.5246 - sparse_categorical_accuracy: 0.9027 - val_loss: 0.6298 - val_sparse_ca
tegorical_accuracy: 0.8674 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.5217 - sparse_categorical_accuracy: 0.9103 - val_loss: 0.5305 - val_sparse_ca
tegorical_accuracy: 0.9024 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.5101 - sparse_categorical_accuracy: 0.9056 - val_loss: 0.5445 - val_sparse_ca
tegorical_accuracy: 0.9037 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.5292 - sparse_categorical_accuracy: 0.9019 - val_loss: 0.5974 - val_sparse_ca
tegorical_accuracy: 0.8994 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.5270 - sparse_categorical_accuracy: 0.9023 - val_loss: 0.5505 - val_sparse_ca
tegorical_accuracy: 0.9038 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.5410 - sparse_categorical_accuracy: 0.9030 - val_loss: 0.6491 - val_sparse_ca
tegorical_accuracy: 0.8897 - 1s/epoch - 3ms/step
Epoch 55/100
469/469 - 1s - loss: 0.5543 - sparse_categorical_accuracy: 0.9009 - val_loss: 0.5779 - val_sparse_ca
tegorical_accuracy: 0.8766 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.5547 - sparse_categorical_accuracy: 0.8972 - val_loss: 0.5389 - val_sparse_ca
tegorical_accuracy: 0.8952 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 0.5374 - sparse_categorical_accuracy: 0.9034 - val_loss: 0.5615 - val_sparse_ca
tegorical_accuracy: 0.8962 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.5468 - sparse_categorical_accuracy: 0.9055 - val_loss: 0.5500 - val_sparse_ca
tegorical_accuracy: 0.8993 - 1s/epoch - 3ms/step
Epoch 59/100
469/469 - 1s - loss: 0.5367 - sparse_categorical_accuracy: 0.9031 - val_loss: 0.5999 - val_sparse_ca
tegorical_accuracy: 0.8659 - 1s/epoch - 3ms/step
Epoch 60/100
469/469 - 1s - loss: 0.5524 - sparse_categorical_accuracy: 0.8960 - val_loss: 0.6742 - val_sparse_ca
tegorical_accuracy: 0.8928 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.5751 - sparse_categorical_accuracy: 0.8956 - val_loss: 0.6086 - val_sparse_ca
tegorical_accuracy: 0.8851 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.5808 - sparse_categorical_accuracy: 0.8928 - val_loss: 0.5821 - val_sparse_ca
tegorical_accuracy: 0.8905 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 0.5736 - sparse_categorical_accuracy: 0.8953 - val_loss: 0.7107 - val_sparse_ca
tegorical_accuracy: 0.8543 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.5782 - sparse_categorical_accuracy: 0.8914 - val_loss: 0.5924 - val_sparse_ca
tegorical_accuracy: 0.8828 - 1s/epoch - 3ms/step
Epoch 65/100
469/469 - 2s - loss: 0.5865 - sparse_categorical_accuracy: 0.8967 - val_loss: 0.7893 - val_sparse_ca
tegorical_accuracy: 0.8721 - 2s/epoch - 4ms/step
Epoch 66/100
469/469 - 2s - loss: 0.6050 - sparse_categorical_accuracy: 0.8927 - val_loss: 0.6272 - val_sparse_ca
tegorical_accuracy: 0.9009 - 2s/epoch - 3ms/step
Epoch 67/100
469/469 - 2s - loss: 0.5822 - sparse_categorical_accuracy: 0.8936 - val_loss: 0.6235 - val_sparse_ca
tegorical_accuracy: 0.8843 - 2s/epoch - 4ms/step
Epoch 68/100
469/469 - 1s - loss: 0.5931 - sparse_categorical_accuracy: 0.8951 - val_loss: 0.5807 - val_sparse_ca
tegorical_accuracy: 0.9054 - 1s/epoch - 3ms/step
Epoch 69/100
469/469 - 2s - loss: 0.5898 - sparse_categorical_accuracy: 0.8964 - val_loss: 0.6483 - val_sparse_ca
tegorical_accuracy: 0.9004 - 2s/epoch - 3ms/step
Epoch 70/100
469/469 - 1s - loss: 0.6419 - sparse_categorical_accuracy: 0.8877 - val_loss: 0.8383 - val_sparse_ca
tegorical_accuracy: 0.8705 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.5759 - sparse_categorical_accuracy: 0.8971 - val_loss: 0.6148 - val_sparse_ca
tegorical_accuracy: 0.8750 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.5814 - sparse_categorical_accuracy: 0.8969 - val_loss: 0.6110 - val_sparse_ca
tegorical_accuracy: 0.8949 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 2s - loss: 0.5818 - sparse_categorical_accuracy: 0.8948 - val_loss: 0.6313 - val_sparse_ca
tegorical_accuracy: 0.8644 - 2s/epoch - 3ms/step
Epoch 74/100
469/469 - 1s - loss: 0.6087 - sparse_categorical_accuracy: 0.8879 - val_loss: 0.6151 - val_sparse_ca
tegorical_accuracy: 0.8707 - 1s/epoch - 2ms/step
```

```
Epoch 75/100
469/469 - 1s - loss: 0.5996 - sparse_categorical_accuracy: 0.8891 - val_loss: 0.6280 - val_sparse_ca
tegorical_accuracy: 0.8620 - 1s/epoch - 3ms/step
Epoch 76/100
469/469 - 1s - loss: 0.6199 - sparse_categorical_accuracy: 0.8858 - val_loss: 0.6651 - val_sparse_ca
tegorical_accuracy: 0.8880 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 2s - loss: 0.6134 - sparse_categorical_accuracy: 0.8888 - val_loss: 0.6471 - val_sparse_ca
tegorical_accuracy: 0.8847 - 2s/epoch - 3ms/step
Epoch 78/100
469/469 - 1s - loss: 0.6424 - sparse_categorical_accuracy: 0.8856 - val_loss: 0.6666 - val_sparse_ca
tegorical_accuracy: 0.8835 - 1s/epoch - 3ms/step
Epoch 79/100
469/469 - 1s - loss: 0.6243 - sparse_categorical_accuracy: 0.8808 - val_loss: 0.6365 - val_sparse_ca
tegorical_accuracy: 0.8833 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.6350 - sparse_categorical_accuracy: 0.8816 - val_loss: 0.6672 - val_sparse_ca
tegorical_accuracy: 0.8966 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.6122 - sparse_categorical_accuracy: 0.8867 - val_loss: 0.6909 - val_sparse_ca
tegorical_accuracy: 0.8643 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.6354 - sparse_categorical_accuracy: 0.8843 - val_loss: 0.6856 - val_sparse_ca
tegorical_accuracy: 0.8839 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.6677 - sparse_categorical_accuracy: 0.8698 - val_loss: 0.6521 - val_sparse_ca
tegorical_accuracy: 0.8739 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.6601 - sparse_categorical_accuracy: 0.8777 - val_loss: 0.7349 - val_sparse_ca
tegorical_accuracy: 0.8861 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.6471 - sparse_categorical_accuracy: 0.8862 - val_loss: 0.7369 - val_sparse_ca
tegorical_accuracy: 0.8768 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.6502 - sparse_categorical_accuracy: 0.8759 - val_loss: 0.6959 - val_sparse_ca
tegorical_accuracy: 0.8589 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.6865 - sparse_categorical_accuracy: 0.8784 - val_loss: 0.7060 - val_sparse_ca
tegorical_accuracy: 0.8768 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.6689 - sparse_categorical_accuracy: 0.8803 - val_loss: 0.6196 - val_sparse_ca
tegorical_accuracy: 0.8943 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.6412 - sparse_categorical_accuracy: 0.8887 - val_loss: 0.6634 - val_sparse_ca
tegorical_accuracy: 0.8734 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.6301 - sparse_categorical_accuracy: 0.8729 - val_loss: 0.7509 - val_sparse_ca
tegorical_accuracy: 0.8599 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.7028 - sparse_categorical_accuracy: 0.8662 - val_loss: 0.7239 - val_sparse_ca
tegorical_accuracy: 0.8449 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.6902 - sparse_categorical_accuracy: 0.8672 - val_loss: 0.7597 - val_sparse_ca
tegorical_accuracy: 0.8224 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.6803 - sparse_categorical_accuracy: 0.8790 - val_loss: 0.6961 - val_sparse_ca
tegorical_accuracy: 0.8835 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.6699 - sparse_categorical_accuracy: 0.8742 - val_loss: 0.7243 - val_sparse_ca
tegorical_accuracy: 0.8930 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.7239 - sparse_categorical_accuracy: 0.8764 - val_loss: 0.6776 - val_sparse_ca
tegorical_accuracy: 0.8711 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.6786 - sparse_categorical_accuracy: 0.8727 - val_loss: 0.7612 - val_sparse_ca
tegorical_accuracy: 0.8493 - 1s/epoch - 3ms/step
Epoch 97/100
469/469 - 1s - loss: 0.7037 - sparse_categorical_accuracy: 0.8675 - val_loss: 0.7619 - val_sparse_ca
tegorical_accuracy: 0.8855 - 1s/epoch - 3ms/step
Epoch 98/100
469/469 - 1s - loss: 0.6866 - sparse_categorical_accuracy: 0.8696 - val_loss: 0.7120 - val_sparse_ca
tegorical_accuracy: 0.8720 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.7141 - sparse_categorical_accuracy: 0.8739 - val_loss: 0.7131 - val_sparse_ca
tegorical_accuracy: 0.8701 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.6803 - sparse_categorical_accuracy: 0.8779 - val_loss: 0.6974 - val_sparse_ca
tegorical_accuracy: 0.8802 - 1s/epoch - 2ms/step
```

In [53]:

```
model6.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.6974 - sparse_categorical_accuracy: 0.8802 - 403ms/epoch - 1ms/step
```

Out[53]:

```
[0.6973861455917358, 0.8802000284194946]
```

Model7: Regularization penalty is L1 with parameter 10^{-3} combined with L2 with parameter 10^{-5}

In [55]:

```
model7 = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform',
  activity_regularizer = tf.keras.regularizers.L1L2(l1=0.001,l2=0.00001)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
```

In [56]:

```
model7.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [57]:

```
history7 = model7.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 2s - loss: 1.4858 - sparse_categorical_accuracy: 0.4694 - val_loss: 1.1487 - val_sparse_ca
tegorical_accuracy: 0.6364 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 1.0108 - sparse_categorical_accuracy: 0.6786 - val_loss: 0.9034 - val_sparse_ca
tegorical_accuracy: 0.7598 - 1s/epoch - 3ms/step
Epoch 3/100
469/469 - 1s - loss: 0.7963 - sparse_categorical_accuracy: 0.8011 - val_loss: 0.7476 - val_sparse_ca
tegorical_accuracy: 0.8258 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 0.7181 - sparse_categorical_accuracy: 0.8376 - val_loss: 0.7191 - val_sparse_ca
tegorical_accuracy: 0.8403 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 0.6912 - sparse_categorical_accuracy: 0.8443 - val_loss: 0.8240 - val_sparse_ca
tegorical_accuracy: 0.8083 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 0.6826 - sparse_categorical_accuracy: 0.8474 - val_loss: 0.6843 - val_sparse_ca
tegorical_accuracy: 0.8511 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.6780 - sparse_categorical_accuracy: 0.8486 - val_loss: 0.7132 - val_sparse_ca
tegorical_accuracy: 0.8531 - 1s/epoch - 3ms/step
Epoch 8/100
469/469 - 1s - loss: 0.6750 - sparse_categorical_accuracy: 0.8492 - val_loss: 0.7213 - val_sparse_ca
tegorical_accuracy: 0.8471 - 1s/epoch - 3ms/step
Epoch 9/100
469/469 - 1s - loss: 0.6666 - sparse_categorical_accuracy: 0.8528 - val_loss: 0.6715 - val_sparse_ca
tegorical_accuracy: 0.8553 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 0.6839 - sparse_categorical_accuracy: 0.8502 - val_loss: 0.7028 - val_sparse_ca
tegorical_accuracy: 0.8476 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 0.6751 - sparse_categorical_accuracy: 0.8515 - val_loss: 0.6929 - val_sparse_ca
tegorical_accuracy: 0.8550 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.6710 - sparse_categorical_accuracy: 0.8538 - val_loss: 0.6756 - val_sparse_ca
tegorical_accuracy: 0.8598 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.6815 - sparse_categorical_accuracy: 0.8514 - val_loss: 0.7014 - val_sparse_ca
tegorical_accuracy: 0.8511 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.6833 - sparse_categorical_accuracy: 0.8523 - val_loss: 0.6578 - val_sparse_ca
tegorical_accuracy: 0.8624 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 0.6841 - sparse_categorical_accuracy: 0.8511 - val_loss: 0.7043 - val_sparse_ca
```

```
tegorical_accuracy: 0.8532 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.7030 - sparse_categorical_accuracy: 0.8464 - val_loss: 0.6866 - val_sparse_ca
tegorical_accuracy: 0.8578 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.7077 - sparse_categorical_accuracy: 0.8474 - val_loss: 0.7863 - val_sparse_ca
tegorical_accuracy: 0.8479 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 2s - loss: 0.7045 - sparse_categorical_accuracy: 0.8480 - val_loss: 0.7143 - val_sparse_ca
tegorical_accuracy: 0.8540 - 2s/epoch - 3ms/step
Epoch 19/100
469/469 - 1s - loss: 0.6959 - sparse_categorical_accuracy: 0.8507 - val_loss: 0.7235 - val_sparse_ca
tegorical_accuracy: 0.8519 - 1s/epoch - 3ms/step
Epoch 20/100
469/469 - 1s - loss: 0.7150 - sparse_categorical_accuracy: 0.8452 - val_loss: 0.7111 - val_sparse_ca
tegorical_accuracy: 0.8450 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.7268 - sparse_categorical_accuracy: 0.8450 - val_loss: 0.7608 - val_sparse_ca
tegorical_accuracy: 0.8464 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.7201 - sparse_categorical_accuracy: 0.8486 - val_loss: 0.7299 - val_sparse_ca
tegorical_accuracy: 0.8426 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.7033 - sparse_categorical_accuracy: 0.8551 - val_loss: 0.8167 - val_sparse_ca
tegorical_accuracy: 0.8227 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.7343 - sparse_categorical_accuracy: 0.8452 - val_loss: 0.7835 - val_sparse_ca
tegorical_accuracy: 0.8420 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.7560 - sparse_categorical_accuracy: 0.8428 - val_loss: 0.8074 - val_sparse_ca
tegorical_accuracy: 0.8200 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.7392 - sparse_categorical_accuracy: 0.8462 - val_loss: 0.8120 - val_sparse_ca
tegorical_accuracy: 0.8109 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.7385 - sparse_categorical_accuracy: 0.8441 - val_loss: 0.6904 - val_sparse_ca
tegorical_accuracy: 0.8581 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.7532 - sparse_categorical_accuracy: 0.8419 - val_loss: 0.7314 - val_sparse_ca
tegorical_accuracy: 0.8509 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.7538 - sparse_categorical_accuracy: 0.8423 - val_loss: 0.7598 - val_sparse_ca
tegorical_accuracy: 0.8400 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.7645 - sparse_categorical_accuracy: 0.8442 - val_loss: 0.7247 - val_sparse_ca
tegorical_accuracy: 0.8560 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.7850 - sparse_categorical_accuracy: 0.8396 - val_loss: 0.7740 - val_sparse_ca
tegorical_accuracy: 0.8419 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.7676 - sparse_categorical_accuracy: 0.8415 - val_loss: 0.8112 - val_sparse_ca
tegorical_accuracy: 0.8452 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.7858 - sparse_categorical_accuracy: 0.8382 - val_loss: 0.7987 - val_sparse_ca
tegorical_accuracy: 0.8437 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.8017 - sparse_categorical_accuracy: 0.8356 - val_loss: 0.8577 - val_sparse_ca
tegorical_accuracy: 0.8147 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.8090 - sparse_categorical_accuracy: 0.8326 - val_loss: 0.8249 - val_sparse_ca
tegorical_accuracy: 0.8266 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.8197 - sparse_categorical_accuracy: 0.8343 - val_loss: 0.8295 - val_sparse_ca
tegorical_accuracy: 0.8306 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.8204 - sparse_categorical_accuracy: 0.8291 - val_loss: 0.8089 - val_sparse_ca
tegorical_accuracy: 0.8481 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.8232 - sparse_categorical_accuracy: 0.8298 - val_loss: 0.8013 - val_sparse_ca
tegorical_accuracy: 0.8311 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.8303 - sparse_categorical_accuracy: 0.8365 - val_loss: 0.8118 - val_sparse_ca
tegorical_accuracy: 0.8268 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 0.8285 - sparse_categorical_accuracy: 0.8346 - val_loss: 0.8778 - val_sparse_ca
tegorical_accuracy: 0.8291 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.8241 - sparse_categorical_accuracy: 0.8345 - val_loss: 0.9649 - val_sparse_ca
tegorical_accuracy: 0.8132 - 1s/epoch - 3ms/step
Epoch 42/100
469/469 - 1s - loss: 0.8382 - sparse_categorical_accuracy: 0.8303 - val_loss: 0.8527 - val_sparse_ca
tegorical_accuracy: 0.8408 - 1s/epoch - 2ms/step
Epoch 43/100
```

```
469/469 - 1s - loss: 0.8398 - sparse_categorical_accuracy: 0.8272 - val_loss: 0.8884 - val_sparse_ca
tegorical_accuracy: 0.8265 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.8142 - sparse_categorical_accuracy: 0.8334 - val_loss: 0.8378 - val_sparse_ca
tegorical_accuracy: 0.8304 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.8432 - sparse_categorical_accuracy: 0.8260 - val_loss: 0.8314 - val_sparse_ca
tegorical_accuracy: 0.8301 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.8510 - sparse_categorical_accuracy: 0.8176 - val_loss: 0.8759 - val_sparse_ca
tegorical_accuracy: 0.8113 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.8463 - sparse_categorical_accuracy: 0.8255 - val_loss: 0.8916 - val_sparse_ca
tegorical_accuracy: 0.8295 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.8556 - sparse_categorical_accuracy: 0.8251 - val_loss: 0.8575 - val_sparse_ca
tegorical_accuracy: 0.8352 - 1s/epoch - 3ms/step
Epoch 49/100
469/469 - 2s - loss: 0.8703 - sparse_categorical_accuracy: 0.8135 - val_loss: 0.8711 - val_sparse_ca
tegorical_accuracy: 0.7995 - 2s/epoch - 4ms/step
Epoch 50/100
469/469 - 2s - loss: 0.9180 - sparse_categorical_accuracy: 0.8141 - val_loss: 0.8704 - val_sparse_ca
tegorical_accuracy: 0.8406 - 2s/epoch - 4ms/step
Epoch 51/100
469/469 - 2s - loss: 0.8453 - sparse_categorical_accuracy: 0.8273 - val_loss: 0.9165 - val_sparse_ca
tegorical_accuracy: 0.8343 - 2s/epoch - 3ms/step
Epoch 52/100
469/469 - 1s - loss: 0.8619 - sparse_categorical_accuracy: 0.8278 - val_loss: 0.9396 - val_sparse_ca
tegorical_accuracy: 0.8154 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.8844 - sparse_categorical_accuracy: 0.8223 - val_loss: 0.9577 - val_sparse_ca
tegorical_accuracy: 0.8324 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.8799 - sparse_categorical_accuracy: 0.8132 - val_loss: 0.8728 - val_sparse_ca
tegorical_accuracy: 0.8040 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.8819 - sparse_categorical_accuracy: 0.8168 - val_loss: 0.8766 - val_sparse_ca
tegorical_accuracy: 0.8261 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.9141 - sparse_categorical_accuracy: 0.8116 - val_loss: 0.9040 - val_sparse_ca
tegorical_accuracy: 0.8101 - 1s/epoch - 3ms/step
Epoch 57/100
469/469 - 1s - loss: 0.9154 - sparse_categorical_accuracy: 0.8095 - val_loss: 0.8322 - val_sparse_ca
tegorical_accuracy: 0.8394 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.9041 - sparse_categorical_accuracy: 0.8047 - val_loss: 0.8983 - val_sparse_ca
tegorical_accuracy: 0.8009 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.9158 - sparse_categorical_accuracy: 0.8130 - val_loss: 0.9161 - val_sparse_ca
tegorical_accuracy: 0.8379 - 1s/epoch - 3ms/step
Epoch 60/100
469/469 - 1s - loss: 0.9049 - sparse_categorical_accuracy: 0.8088 - val_loss: 0.9106 - val_sparse_ca
tegorical_accuracy: 0.8054 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.9312 - sparse_categorical_accuracy: 0.8019 - val_loss: 0.9229 - val_sparse_ca
tegorical_accuracy: 0.7819 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.9062 - sparse_categorical_accuracy: 0.8006 - val_loss: 1.1287 - val_sparse_ca
tegorical_accuracy: 0.7924 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 0.9343 - sparse_categorical_accuracy: 0.8001 - val_loss: 0.9415 - val_sparse_ca
tegorical_accuracy: 0.8035 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.9456 - sparse_categorical_accuracy: 0.7999 - val_loss: 1.0569 - val_sparse_ca
tegorical_accuracy: 0.8170 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.9522 - sparse_categorical_accuracy: 0.8195 - val_loss: 0.9242 - val_sparse_ca
tegorical_accuracy: 0.8260 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.9281 - sparse_categorical_accuracy: 0.8137 - val_loss: 0.9548 - val_sparse_ca
tegorical_accuracy: 0.8227 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.9536 - sparse_categorical_accuracy: 0.7986 - val_loss: 0.9710 - val_sparse_ca
tegorical_accuracy: 0.8090 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 0.9400 - sparse_categorical_accuracy: 0.8034 - val_loss: 0.9382 - val_sparse_ca
tegorical_accuracy: 0.8115 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 0.9614 - sparse_categorical_accuracy: 0.8037 - val_loss: 1.0645 - val_sparse_ca
tegorical_accuracy: 0.8133 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.9472 - sparse_categorical_accuracy: 0.7864 - val_loss: 0.9302 - val_sparse_ca
tegorical_accuracy: 0.7815 - 1s/epoch - 2ms/step
```

```
Epoch 71/100
469/469 - 1s - loss: 0.9419 - sparse_categorical_accuracy: 0.7857 - val_loss: 0.9109 - val_sparse_ca
tegorical_accuracy: 0.8099 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.9550 - sparse_categorical_accuracy: 0.7852 - val_loss: 1.0042 - val_sparse_ca
tegorical_accuracy: 0.7726 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.9714 - sparse_categorical_accuracy: 0.7767 - val_loss: 0.9657 - val_sparse_ca
tegorical_accuracy: 0.7703 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.9581 - sparse_categorical_accuracy: 0.7779 - val_loss: 0.9677 - val_sparse_ca
tegorical_accuracy: 0.7964 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 1.0091 - sparse_categorical_accuracy: 0.7806 - val_loss: 0.9799 - val_sparse_ca
tegorical_accuracy: 0.7730 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.9983 - sparse_categorical_accuracy: 0.7908 - val_loss: 0.9620 - val_sparse_ca
tegorical_accuracy: 0.8020 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.9634 - sparse_categorical_accuracy: 0.8008 - val_loss: 1.0824 - val_sparse_ca
tegorical_accuracy: 0.8016 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.9699 - sparse_categorical_accuracy: 0.7774 - val_loss: 1.0033 - val_sparse_ca
tegorical_accuracy: 0.7839 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.9708 - sparse_categorical_accuracy: 0.7839 - val_loss: 0.9623 - val_sparse_ca
tegorical_accuracy: 0.7883 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 1.0014 - sparse_categorical_accuracy: 0.7918 - val_loss: 0.9724 - val_sparse_ca
tegorical_accuracy: 0.8141 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.9525 - sparse_categorical_accuracy: 0.7975 - val_loss: 1.0260 - val_sparse_ca
tegorical_accuracy: 0.7669 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.9543 - sparse_categorical_accuracy: 0.8042 - val_loss: 0.9636 - val_sparse_ca
tegorical_accuracy: 0.8222 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.9761 - sparse_categorical_accuracy: 0.7950 - val_loss: 0.9910 - val_sparse_ca
tegorical_accuracy: 0.8103 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.9837 - sparse_categorical_accuracy: 0.8044 - val_loss: 1.0178 - val_sparse_ca
tegorical_accuracy: 0.8073 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.9918 - sparse_categorical_accuracy: 0.8019 - val_loss: 0.9721 - val_sparse_ca
tegorical_accuracy: 0.7787 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 1.0053 - sparse_categorical_accuracy: 0.7844 - val_loss: 0.9263 - val_sparse_ca
tegorical_accuracy: 0.8034 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.9882 - sparse_categorical_accuracy: 0.7954 - val_loss: 0.9619 - val_sparse_ca
tegorical_accuracy: 0.7698 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.9951 - sparse_categorical_accuracy: 0.7839 - val_loss: 1.0013 - val_sparse_ca
tegorical_accuracy: 0.8013 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.9932 - sparse_categorical_accuracy: 0.7849 - val_loss: 1.0043 - val_sparse_ca
tegorical_accuracy: 0.8362 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 1.0006 - sparse_categorical_accuracy: 0.7720 - val_loss: 1.0000 - val_sparse_ca
tegorical_accuracy: 0.8140 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 1.0008 - sparse_categorical_accuracy: 0.7724 - val_loss: 0.9936 - val_sparse_ca
tegorical_accuracy: 0.7554 - 1s/epoch - 3ms/step
Epoch 92/100
469/469 - 1s - loss: 1.0191 - sparse_categorical_accuracy: 0.7767 - val_loss: 1.0584 - val_sparse_ca
tegorical_accuracy: 0.7247 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 1.0292 - sparse_categorical_accuracy: 0.7678 - val_loss: 1.1908 - val_sparse_ca
tegorical_accuracy: 0.7577 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 1.0447 - sparse_categorical_accuracy: 0.7605 - val_loss: 1.0402 - val_sparse_ca
tegorical_accuracy: 0.7476 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 1.0702 - sparse_categorical_accuracy: 0.7608 - val_loss: 0.9962 - val_sparse_ca
tegorical_accuracy: 0.7748 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 1.0610 - sparse_categorical_accuracy: 0.7669 - val_loss: 0.9887 - val_sparse_ca
tegorical_accuracy: 0.7887 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 1.0193 - sparse_categorical_accuracy: 0.7877 - val_loss: 1.1105 - val_sparse_ca
tegorical_accuracy: 0.7790 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 1.0775 - sparse_categorical_accuracy: 0.7626 - val_loss: 1.1312 - val_sparse_ca
```

```
tegorical_accuracy: 0.7654 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 1.0702 - sparse_categorical_accuracy: 0.7745 - val_loss: 1.1295 - val_sparse_ca
tegorical_accuracy: 0.6839 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 1.0517 - sparse_categorical_accuracy: 0.7660 - val_loss: 0.9829 - val_sparse_ca
tegorical_accuracy: 0.7897 - 1s/epoch - 2ms/step
```

In [58]:

```
model7.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.9829 - sparse_categorical_accuracy: 0.7897 - 378ms/epoch - 1ms/step
```

Out[58]:

```
[0.9829106330871582, 0.7896999716758728]
```

Model8: Dropout with parameter 0.1

In [83]:

```
model8 = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dropout(0.1),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
model8.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
history8 = model8.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 2s - loss: 0.5109 - sparse_categorical_accuracy: 0.8400 - val_loss: 0.2626 - val_sparse_ca
tegorical_accuracy: 0.9221 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 0.3253 - sparse_categorical_accuracy: 0.9011 - val_loss: 0.2168 - val_sparse_ca
tegorical_accuracy: 0.9343 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 0.2922 - sparse_categorical_accuracy: 0.9110 - val_loss: 0.1969 - val_sparse_ca
tegorical_accuracy: 0.9427 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 0.2685 - sparse_categorical_accuracy: 0.9182 - val_loss: 0.2039 - val_sparse_ca
tegorical_accuracy: 0.9387 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 0.2565 - sparse_categorical_accuracy: 0.9217 - val_loss: 0.1927 - val_sparse_ca
tegorical_accuracy: 0.9428 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 0.2491 - sparse_categorical_accuracy: 0.9240 - val_loss: 0.1946 - val_sparse_ca
tegorical_accuracy: 0.9430 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.2395 - sparse_categorical_accuracy: 0.9263 - val_loss: 0.2014 - val_sparse_ca
tegorical_accuracy: 0.9397 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 0.2334 - sparse_categorical_accuracy: 0.9276 - val_loss: 0.1831 - val_sparse_ca
tegorical_accuracy: 0.9453 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 0.2267 - sparse_categorical_accuracy: 0.9305 - val_loss: 0.1708 - val_sparse_ca
tegorical_accuracy: 0.9492 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 0.2209 - sparse_categorical_accuracy: 0.9323 - val_loss: 0.1657 - val_sparse_ca
tegorical_accuracy: 0.9513 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 0.2190 - sparse_categorical_accuracy: 0.9322 - val_loss: 0.1686 - val_sparse_ca
tegorical_accuracy: 0.9522 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.2156 - sparse_categorical_accuracy: 0.9341 - val_loss: 0.1634 - val_sparse_ca
tegorical_accuracy: 0.9528 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.2144 - sparse_categorical_accuracy: 0.9342 - val_loss: 0.1636 - val_sparse_ca
tegorical_accuracy: 0.9517 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.2099 - sparse_categorical_accuracy: 0.9357 - val_loss: 0.1630 - val_sparse_ca
tegorical_accuracy: 0.9518 - 1s/epoch - 2ms/step
Epoch 15/100
```

```
469/469 - 1s - loss: 0.2095 - sparse_categorical_accuracy: 0.9342 - val_loss: 0.1656 - val_sparse_ca
tegorical_accuracy: 0.9532 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.2033 - sparse_categorical_accuracy: 0.9374 - val_loss: 0.1746 - val_sparse_ca
tegorical_accuracy: 0.9493 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.2077 - sparse_categorical_accuracy: 0.9364 - val_loss: 0.1629 - val_sparse_ca
tegorical_accuracy: 0.9524 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 0.2018 - sparse_categorical_accuracy: 0.9367 - val_loss: 0.1640 - val_sparse_ca
tegorical_accuracy: 0.9525 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.2040 - sparse_categorical_accuracy: 0.9368 - val_loss: 0.1772 - val_sparse_ca
tegorical_accuracy: 0.9500 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.1992 - sparse_categorical_accuracy: 0.9388 - val_loss: 0.1597 - val_sparse_ca
tegorical_accuracy: 0.9530 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.1954 - sparse_categorical_accuracy: 0.9384 - val_loss: 0.1616 - val_sparse_ca
tegorical_accuracy: 0.9542 - 1s/epoch - 3ms/step
Epoch 22/100
469/469 - 2s - loss: 0.1973 - sparse_categorical_accuracy: 0.9394 - val_loss: 0.1638 - val_sparse_ca
tegorical_accuracy: 0.9528 - 2s/epoch - 3ms/step
Epoch 23/100
469/469 - 1s - loss: 0.1963 - sparse_categorical_accuracy: 0.9392 - val_loss: 0.1691 - val_sparse_ca
tegorical_accuracy: 0.9515 - 1s/epoch - 3ms/step
Epoch 24/100
469/469 - 2s - loss: 0.1970 - sparse_categorical_accuracy: 0.9386 - val_loss: 0.1777 - val_sparse_ca
tegorical_accuracy: 0.9496 - 2s/epoch - 3ms/step
Epoch 25/100
469/469 - 2s - loss: 0.1930 - sparse_categorical_accuracy: 0.9405 - val_loss: 0.1744 - val_sparse_ca
tegorical_accuracy: 0.9506 - 2s/epoch - 3ms/step
Epoch 26/100
469/469 - 1s - loss: 0.1929 - sparse_categorical_accuracy: 0.9399 - val_loss: 0.1695 - val_sparse_ca
tegorical_accuracy: 0.9487 - 1s/epoch - 3ms/step
Epoch 27/100
469/469 - 1s - loss: 0.1956 - sparse_categorical_accuracy: 0.9391 - val_loss: 0.1651 - val_sparse_ca
tegorical_accuracy: 0.9534 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.1891 - sparse_categorical_accuracy: 0.9416 - val_loss: 0.1751 - val_sparse_ca
tegorical_accuracy: 0.9510 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.1928 - sparse_categorical_accuracy: 0.9408 - val_loss: 0.1704 - val_sparse_ca
tegorical_accuracy: 0.9512 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.1963 - sparse_categorical_accuracy: 0.9390 - val_loss: 0.1628 - val_sparse_ca
tegorical_accuracy: 0.9542 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.1903 - sparse_categorical_accuracy: 0.9411 - val_loss: 0.1730 - val_sparse_ca
tegorical_accuracy: 0.9502 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 2s - loss: 0.1887 - sparse_categorical_accuracy: 0.9407 - val_loss: 0.1727 - val_sparse_ca
tegorical_accuracy: 0.9512 - 2s/epoch - 3ms/step
Epoch 33/100
469/469 - 1s - loss: 0.1846 - sparse_categorical_accuracy: 0.9430 - val_loss: 0.1643 - val_sparse_ca
tegorical_accuracy: 0.9538 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.1865 - sparse_categorical_accuracy: 0.9414 - val_loss: 0.1712 - val_sparse_ca
tegorical_accuracy: 0.9522 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.1873 - sparse_categorical_accuracy: 0.9416 - val_loss: 0.1726 - val_sparse_ca
tegorical_accuracy: 0.9519 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.1876 - sparse_categorical_accuracy: 0.9424 - val_loss: 0.1689 - val_sparse_ca
tegorical_accuracy: 0.9548 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.1850 - sparse_categorical_accuracy: 0.9426 - val_loss: 0.1801 - val_sparse_ca
tegorical_accuracy: 0.9485 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.1849 - sparse_categorical_accuracy: 0.9432 - val_loss: 0.1722 - val_sparse_ca
tegorical_accuracy: 0.9516 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.1867 - sparse_categorical_accuracy: 0.9409 - val_loss: 0.1769 - val_sparse_ca
tegorical_accuracy: 0.9522 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 0.1847 - sparse_categorical_accuracy: 0.9423 - val_loss: 0.1791 - val_sparse_ca
tegorical_accuracy: 0.9511 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.1848 - sparse_categorical_accuracy: 0.9426 - val_loss: 0.1784 - val_sparse_ca
tegorical_accuracy: 0.9499 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.1814 - sparse_categorical_accuracy: 0.9434 - val_loss: 0.1753 - val_sparse_ca
tegorical_accuracy: 0.9516 - 1s/epoch - 2ms/step
```

```
Epoch 43/100
469/469 - 1s - loss: 0.1821 - sparse_categorical_accuracy: 0.9433 - val_loss: 0.1802 - val_sparse_ca
tegorical_accuracy: 0.9498 - 1s/epoch - 3ms/step
Epoch 44/100
469/469 - 1s - loss: 0.1797 - sparse_categorical_accuracy: 0.9449 - val_loss: 0.1693 - val_sparse_ca
tegorical_accuracy: 0.9514 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.1847 - sparse_categorical_accuracy: 0.9427 - val_loss: 0.1745 - val_sparse_ca
tegorical_accuracy: 0.9523 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.1865 - sparse_categorical_accuracy: 0.9418 - val_loss: 0.1685 - val_sparse_ca
tegorical_accuracy: 0.9528 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.1838 - sparse_categorical_accuracy: 0.9428 - val_loss: 0.1692 - val_sparse_ca
tegorical_accuracy: 0.9525 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.1797 - sparse_categorical_accuracy: 0.9446 - val_loss: 0.1733 - val_sparse_ca
tegorical_accuracy: 0.9522 - 1s/epoch - 3ms/step
Epoch 49/100
469/469 - 1s - loss: 0.1797 - sparse_categorical_accuracy: 0.9441 - val_loss: 0.1821 - val_sparse_ca
tegorical_accuracy: 0.9511 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.1806 - sparse_categorical_accuracy: 0.9438 - val_loss: 0.1753 - val_sparse_ca
tegorical_accuracy: 0.9506 - 1s/epoch - 3ms/step
Epoch 51/100
469/469 - 1s - loss: 0.1783 - sparse_categorical_accuracy: 0.9452 - val_loss: 0.1875 - val_sparse_ca
tegorical_accuracy: 0.9498 - 1s/epoch - 3ms/step
Epoch 52/100
469/469 - 1s - loss: 0.1854 - sparse_categorical_accuracy: 0.9424 - val_loss: 0.1760 - val_sparse_ca
tegorical_accuracy: 0.9506 - 1s/epoch - 3ms/step
Epoch 53/100
469/469 - 2s - loss: 0.1797 - sparse_categorical_accuracy: 0.9446 - val_loss: 0.1809 - val_sparse_ca
tegorical_accuracy: 0.9479 - 2s/epoch - 3ms/step
Epoch 54/100
469/469 - 1s - loss: 0.1791 - sparse_categorical_accuracy: 0.9443 - val_loss: 0.1836 - val_sparse_ca
tegorical_accuracy: 0.9491 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.1773 - sparse_categorical_accuracy: 0.9449 - val_loss: 0.1734 - val_sparse_ca
tegorical_accuracy: 0.9514 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.1745 - sparse_categorical_accuracy: 0.9445 - val_loss: 0.1851 - val_sparse_ca
tegorical_accuracy: 0.9496 - 1s/epoch - 3ms/step
Epoch 57/100
469/469 - 1s - loss: 0.1737 - sparse_categorical_accuracy: 0.9455 - val_loss: 0.1924 - val_sparse_ca
tegorical_accuracy: 0.9464 - 1s/epoch - 3ms/step
Epoch 58/100
469/469 - 1s - loss: 0.1773 - sparse_categorical_accuracy: 0.9449 - val_loss: 0.1884 - val_sparse_ca
tegorical_accuracy: 0.9474 - 1s/epoch - 3ms/step
Epoch 59/100
469/469 - 2s - loss: 0.1762 - sparse_categorical_accuracy: 0.9447 - val_loss: 0.1773 - val_sparse_ca
tegorical_accuracy: 0.9532 - 2s/epoch - 4ms/step
Epoch 60/100
469/469 - 2s - loss: 0.1731 - sparse_categorical_accuracy: 0.9463 - val_loss: 0.1709 - val_sparse_ca
tegorical_accuracy: 0.9527 - 2s/epoch - 3ms/step
Epoch 61/100
469/469 - 1s - loss: 0.1764 - sparse_categorical_accuracy: 0.9449 - val_loss: 0.1872 - val_sparse_ca
tegorical_accuracy: 0.9479 - 1s/epoch - 3ms/step
Epoch 62/100
469/469 - 2s - loss: 0.1758 - sparse_categorical_accuracy: 0.9448 - val_loss: 0.1799 - val_sparse_ca
tegorical_accuracy: 0.9510 - 2s/epoch - 4ms/step
Epoch 63/100
469/469 - 2s - loss: 0.1721 - sparse_categorical_accuracy: 0.9465 - val_loss: 0.1741 - val_sparse_ca
tegorical_accuracy: 0.9546 - 2s/epoch - 4ms/step
Epoch 64/100
469/469 - 1s - loss: 0.1790 - sparse_categorical_accuracy: 0.9449 - val_loss: 0.1787 - val_sparse_ca
tegorical_accuracy: 0.9513 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.1731 - sparse_categorical_accuracy: 0.9461 - val_loss: 0.1828 - val_sparse_ca
tegorical_accuracy: 0.9508 - 1s/epoch - 3ms/step
Epoch 66/100
469/469 - 1s - loss: 0.1730 - sparse_categorical_accuracy: 0.9457 - val_loss: 0.1754 - val_sparse_ca
tegorical_accuracy: 0.9503 - 1s/epoch - 3ms/step
Epoch 67/100
469/469 - 2s - loss: 0.1698 - sparse_categorical_accuracy: 0.9470 - val_loss: 0.1771 - val_sparse_ca
tegorical_accuracy: 0.9518 - 2s/epoch - 3ms/step
Epoch 68/100
469/469 - 2s - loss: 0.1746 - sparse_categorical_accuracy: 0.9454 - val_loss: 0.1824 - val_sparse_ca
tegorical_accuracy: 0.9517 - 2s/epoch - 3ms/step
Epoch 69/100
469/469 - 1s - loss: 0.1729 - sparse_categorical_accuracy: 0.9469 - val_loss: 0.1756 - val_sparse_ca
tegorical_accuracy: 0.9531 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.1697 - sparse_categorical_accuracy: 0.9463 - val_loss: 0.1806 - val_sparse_ca
```

```
tegorical_accuracy: 0.9516 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.1725 - sparse_categorical_accuracy: 0.9455 - val_loss: 0.1670 - val_sparse_ca
tegorical_accuracy: 0.9546 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.1698 - sparse_categorical_accuracy: 0.9470 - val_loss: 0.1806 - val_sparse_ca
tegorical_accuracy: 0.9525 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.1710 - sparse_categorical_accuracy: 0.9467 - val_loss: 0.1825 - val_sparse_ca
tegorical_accuracy: 0.9487 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.1710 - sparse_categorical_accuracy: 0.9474 - val_loss: 0.1803 - val_sparse_ca
tegorical_accuracy: 0.9505 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.1683 - sparse_categorical_accuracy: 0.9470 - val_loss: 0.1820 - val_sparse_ca
tegorical_accuracy: 0.9510 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.1735 - sparse_categorical_accuracy: 0.9466 - val_loss: 0.1804 - val_sparse_ca
tegorical_accuracy: 0.9526 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.1711 - sparse_categorical_accuracy: 0.9474 - val_loss: 0.1731 - val_sparse_ca
tegorical_accuracy: 0.9530 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.1672 - sparse_categorical_accuracy: 0.9490 - val_loss: 0.1853 - val_sparse_ca
tegorical_accuracy: 0.9513 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.1698 - sparse_categorical_accuracy: 0.9468 - val_loss: 0.1919 - val_sparse_ca
tegorical_accuracy: 0.9500 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.1723 - sparse_categorical_accuracy: 0.9464 - val_loss: 0.1853 - val_sparse_ca
tegorical_accuracy: 0.9497 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.1737 - sparse_categorical_accuracy: 0.9452 - val_loss: 0.1846 - val_sparse_ca
tegorical_accuracy: 0.9516 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.1672 - sparse_categorical_accuracy: 0.9473 - val_loss: 0.1843 - val_sparse_ca
tegorical_accuracy: 0.9501 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.1670 - sparse_categorical_accuracy: 0.9483 - val_loss: 0.1878 - val_sparse_ca
tegorical_accuracy: 0.9499 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.1693 - sparse_categorical_accuracy: 0.9475 - val_loss: 0.1743 - val_sparse_ca
tegorical_accuracy: 0.9537 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.1713 - sparse_categorical_accuracy: 0.9468 - val_loss: 0.1810 - val_sparse_ca
tegorical_accuracy: 0.9519 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.1690 - sparse_categorical_accuracy: 0.9475 - val_loss: 0.1803 - val_sparse_ca
tegorical_accuracy: 0.9528 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.1683 - sparse_categorical_accuracy: 0.9481 - val_loss: 0.1770 - val_sparse_ca
tegorical_accuracy: 0.9523 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.1703 - sparse_categorical_accuracy: 0.9464 - val_loss: 0.1830 - val_sparse_ca
tegorical_accuracy: 0.9510 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.1670 - sparse_categorical_accuracy: 0.9476 - val_loss: 0.1894 - val_sparse_ca
tegorical_accuracy: 0.9527 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.1709 - sparse_categorical_accuracy: 0.9471 - val_loss: 0.1922 - val_sparse_ca
tegorical_accuracy: 0.9509 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.1650 - sparse_categorical_accuracy: 0.9476 - val_loss: 0.1850 - val_sparse_ca
tegorical_accuracy: 0.9517 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.1695 - sparse_categorical_accuracy: 0.9477 - val_loss: 0.1851 - val_sparse_ca
tegorical_accuracy: 0.9499 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.1670 - sparse_categorical_accuracy: 0.9476 - val_loss: 0.1846 - val_sparse_ca
tegorical_accuracy: 0.9509 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.1662 - sparse_categorical_accuracy: 0.9480 - val_loss: 0.1824 - val_sparse_ca
tegorical_accuracy: 0.9518 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.1642 - sparse_categorical_accuracy: 0.9488 - val_loss: 0.1868 - val_sparse_ca
tegorical_accuracy: 0.9494 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.1652 - sparse_categorical_accuracy: 0.9485 - val_loss: 0.1800 - val_sparse_ca
tegorical_accuracy: 0.9519 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.1676 - sparse_categorical_accuracy: 0.9476 - val_loss: 0.1882 - val_sparse_ca
tegorical_accuracy: 0.9510 - 1s/epoch - 2ms/step
Epoch 98/100
```

```
469/469 - 1s - loss: 0.1687 - sparse_categorical_accuracy: 0.9476 - val_loss: 0.1734 - val_sparse_ca
tegorical_accuracy: 0.9536 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.1649 - sparse_categorical_accuracy: 0.9481 - val_loss: 0.1797 - val_sparse_ca
tegorical_accuracy: 0.9514 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.1661 - sparse_categorical_accuracy: 0.9488 - val_loss: 0.1883 - val_sparse_ca
tegorical_accuracy: 0.9492 - 1s/epoch - 2ms/step
```

In [84]:

```
model8.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.1883 - sparse_categorical_accuracy: 0.9492 - 409ms/epoch - 1ms/step
```

Out[84]:

```
[0.18833673000335693, 0.9491999745368958]
```

Model9: Dropout with parameter 0.2

In [85]:

```
model9 = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dropout(0.2),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
model9.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
history9 = model9.fit(x_train, y_train,
                      batch_size=128,
                      epochs=100,
                      validation_data=(x_test, y_test),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 2s - loss: 0.6270 - sparse_categorical_accuracy: 0.7918 - val_loss: 0.2767 - val_sparse_ca
tegorical_accuracy: 0.9186 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 0.4368 - sparse_categorical_accuracy: 0.8619 - val_loss: 0.2488 - val_sparse_ca
tegorical_accuracy: 0.9252 - 1s/epoch - 3ms/step
Epoch 3/100
469/469 - 1s - loss: 0.3933 - sparse_categorical_accuracy: 0.8758 - val_loss: 0.2437 - val_sparse_ca
tegorical_accuracy: 0.9276 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 0.3786 - sparse_categorical_accuracy: 0.8811 - val_loss: 0.2279 - val_sparse_ca
tegorical_accuracy: 0.9343 - 1s/epoch - 3ms/step
Epoch 5/100
469/469 - 1s - loss: 0.3657 - sparse_categorical_accuracy: 0.8841 - val_loss: 0.2207 - val_sparse_ca
tegorical_accuracy: 0.9331 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 0.3559 - sparse_categorical_accuracy: 0.8881 - val_loss: 0.2186 - val_sparse_ca
tegorical_accuracy: 0.9363 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.3523 - sparse_categorical_accuracy: 0.8878 - val_loss: 0.1996 - val_sparse_ca
tegorical_accuracy: 0.9400 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 0.3457 - sparse_categorical_accuracy: 0.8890 - val_loss: 0.2148 - val_sparse_ca
tegorical_accuracy: 0.9350 - 1s/epoch - 3ms/step
Epoch 9/100
469/469 - 1s - loss: 0.3421 - sparse_categorical_accuracy: 0.8913 - val_loss: 0.2024 - val_sparse_ca
tegorical_accuracy: 0.9394 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 0.3350 - sparse_categorical_accuracy: 0.8944 - val_loss: 0.2172 - val_sparse_ca
tegorical_accuracy: 0.9373 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 0.3312 - sparse_categorical_accuracy: 0.8971 - val_loss: 0.2023 - val_sparse_ca
tegorical_accuracy: 0.9408 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.3277 - sparse_categorical_accuracy: 0.8961 - val_loss: 0.1995 - val_sparse_ca
tegorical_accuracy: 0.9410 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.3257 - sparse_categorical_accuracy: 0.8979 - val_loss: 0.2087 - val_sparse_ca
tegorical_accuracy: 0.9391 - 1s/epoch - 3ms/step
Epoch 14/100
469/469 - 1s - loss: 0.3210 - sparse_categorical_accuracy: 0.8984 - val_loss: 0.1934 - val_sparse_ca
tegorical_accuracy: 0.9425 - 1s/epoch - 2ms/step
```

```
Epoch 15/100
469/469 - 1s - loss: 0.3208 - sparse_categorical_accuracy: 0.8988 - val_loss: 0.1938 - val_sparse_ca
tegorical_accuracy: 0.9438 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.3141 - sparse_categorical_accuracy: 0.9011 - val_loss: 0.2052 - val_sparse_ca
tegorical_accuracy: 0.9407 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.3101 - sparse_categorical_accuracy: 0.9021 - val_loss: 0.1911 - val_sparse_ca
tegorical_accuracy: 0.9450 - 1s/epoch - 3ms/step
Epoch 18/100
469/469 - 1s - loss: 0.3161 - sparse_categorical_accuracy: 0.9020 - val_loss: 0.1990 - val_sparse_ca
tegorical_accuracy: 0.9419 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.3071 - sparse_categorical_accuracy: 0.9039 - val_loss: 0.2107 - val_sparse_ca
tegorical_accuracy: 0.9373 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.3051 - sparse_categorical_accuracy: 0.9051 - val_loss: 0.1911 - val_sparse_ca
tegorical_accuracy: 0.9438 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.3038 - sparse_categorical_accuracy: 0.9044 - val_loss: 0.2050 - val_sparse_ca
tegorical_accuracy: 0.9380 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.3024 - sparse_categorical_accuracy: 0.9064 - val_loss: 0.2029 - val_sparse_ca
tegorical_accuracy: 0.9406 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.2980 - sparse_categorical_accuracy: 0.9070 - val_loss: 0.1877 - val_sparse_ca
tegorical_accuracy: 0.9462 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.2979 - sparse_categorical_accuracy: 0.9076 - val_loss: 0.1892 - val_sparse_ca
tegorical_accuracy: 0.9445 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.2979 - sparse_categorical_accuracy: 0.9088 - val_loss: 0.1909 - val_sparse_ca
tegorical_accuracy: 0.9448 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.2940 - sparse_categorical_accuracy: 0.9078 - val_loss: 0.2080 - val_sparse_ca
tegorical_accuracy: 0.9409 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.2939 - sparse_categorical_accuracy: 0.9095 - val_loss: 0.1919 - val_sparse_ca
tegorical_accuracy: 0.9445 - 1s/epoch - 3ms/step
Epoch 28/100
469/469 - 1s - loss: 0.2925 - sparse_categorical_accuracy: 0.9103 - val_loss: 0.2076 - val_sparse_ca
tegorical_accuracy: 0.9375 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.2887 - sparse_categorical_accuracy: 0.9104 - val_loss: 0.2013 - val_sparse_ca
tegorical_accuracy: 0.9427 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.2906 - sparse_categorical_accuracy: 0.9100 - val_loss: 0.1971 - val_sparse_ca
tegorical_accuracy: 0.9431 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.2902 - sparse_categorical_accuracy: 0.9102 - val_loss: 0.1925 - val_sparse_ca
tegorical_accuracy: 0.9419 - 1s/epoch - 3ms/step
Epoch 32/100
469/469 - 1s - loss: 0.2909 - sparse_categorical_accuracy: 0.9097 - val_loss: 0.1938 - val_sparse_ca
tegorical_accuracy: 0.9438 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.2843 - sparse_categorical_accuracy: 0.9118 - val_loss: 0.1902 - val_sparse_ca
tegorical_accuracy: 0.9455 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.2842 - sparse_categorical_accuracy: 0.9118 - val_loss: 0.1984 - val_sparse_ca
tegorical_accuracy: 0.9426 - 1s/epoch - 3ms/step
Epoch 35/100
469/469 - 1s - loss: 0.2834 - sparse_categorical_accuracy: 0.9130 - val_loss: 0.2018 - val_sparse_ca
tegorical_accuracy: 0.9422 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.2793 - sparse_categorical_accuracy: 0.9137 - val_loss: 0.1963 - val_sparse_ca
tegorical_accuracy: 0.9468 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.2823 - sparse_categorical_accuracy: 0.9136 - val_loss: 0.2077 - val_sparse_ca
tegorical_accuracy: 0.9396 - 1s/epoch - 3ms/step
Epoch 38/100
469/469 - 2s - loss: 0.2774 - sparse_categorical_accuracy: 0.9146 - val_loss: 0.2001 - val_sparse_ca
tegorical_accuracy: 0.9436 - 2s/epoch - 4ms/step
Epoch 39/100
469/469 - 2s - loss: 0.2826 - sparse_categorical_accuracy: 0.9129 - val_loss: 0.1982 - val_sparse_ca
tegorical_accuracy: 0.9443 - 2s/epoch - 4ms/step
Epoch 40/100
469/469 - 2s - loss: 0.2789 - sparse_categorical_accuracy: 0.9141 - val_loss: 0.1988 - val_sparse_ca
tegorical_accuracy: 0.9431 - 2s/epoch - 5ms/step
Epoch 41/100
469/469 - 2s - loss: 0.2780 - sparse_categorical_accuracy: 0.9145 - val_loss: 0.2007 - val_sparse_ca
tegorical_accuracy: 0.9430 - 2s/epoch - 5ms/step
Epoch 42/100
469/469 - 2s - loss: 0.2765 - sparse_categorical_accuracy: 0.9154 - val_loss: 0.2023 - val_sparse_ca
```

```
tegorical_accuracy: 0.9427 - 2s/epoch - 4ms/step
Epoch 43/100
469/469 - 1s - loss: 0.2743 - sparse_categorical_accuracy: 0.9152 - val_loss: 0.2035 - val_sparse_ca
tegorical_accuracy: 0.9426 - 1s/epoch - 3ms/step
Epoch 44/100
469/469 - 1s - loss: 0.2780 - sparse_categorical_accuracy: 0.9137 - val_loss: 0.1887 - val_sparse_ca
tegorical_accuracy: 0.9476 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 2s - loss: 0.2736 - sparse_categorical_accuracy: 0.9147 - val_loss: 0.1961 - val_sparse_ca
tegorical_accuracy: 0.9428 - 2s/epoch - 4ms/step
Epoch 46/100
469/469 - 1s - loss: 0.2696 - sparse_categorical_accuracy: 0.9174 - val_loss: 0.1912 - val_sparse_ca
tegorical_accuracy: 0.9448 - 1s/epoch - 3ms/step
Epoch 47/100
469/469 - 2s - loss: 0.2776 - sparse_categorical_accuracy: 0.9154 - val_loss: 0.2030 - val_sparse_ca
tegorical_accuracy: 0.9417 - 2s/epoch - 4ms/step
Epoch 48/100
469/469 - 2s - loss: 0.2731 - sparse_categorical_accuracy: 0.9154 - val_loss: 0.1934 - val_sparse_ca
tegorical_accuracy: 0.9473 - 2s/epoch - 4ms/step
Epoch 49/100
469/469 - 2s - loss: 0.2651 - sparse_categorical_accuracy: 0.9191 - val_loss: 0.2100 - val_sparse_ca
tegorical_accuracy: 0.9381 - 2s/epoch - 3ms/step
Epoch 50/100
469/469 - 2s - loss: 0.2745 - sparse_categorical_accuracy: 0.9157 - val_loss: 0.1866 - val_sparse_ca
tegorical_accuracy: 0.9478 - 2s/epoch - 3ms/step
Epoch 51/100
469/469 - 2s - loss: 0.2697 - sparse_categorical_accuracy: 0.9176 - val_loss: 0.1823 - val_sparse_ca
tegorical_accuracy: 0.9485 - 2s/epoch - 4ms/step
Epoch 52/100
469/469 - 2s - loss: 0.2732 - sparse_categorical_accuracy: 0.9166 - val_loss: 0.1981 - val_sparse_ca
tegorical_accuracy: 0.9424 - 2s/epoch - 4ms/step
Epoch 53/100
469/469 - 1s - loss: 0.2686 - sparse_categorical_accuracy: 0.9182 - val_loss: 0.1995 - val_sparse_ca
tegorical_accuracy: 0.9449 - 1s/epoch - 3ms/step
Epoch 54/100
469/469 - 2s - loss: 0.2674 - sparse_categorical_accuracy: 0.9176 - val_loss: 0.1909 - val_sparse_ca
tegorical_accuracy: 0.9453 - 2s/epoch - 3ms/step
Epoch 55/100
469/469 - 2s - loss: 0.2658 - sparse_categorical_accuracy: 0.9174 - val_loss: 0.1909 - val_sparse_ca
tegorical_accuracy: 0.9459 - 2s/epoch - 3ms/step
Epoch 56/100
469/469 - 2s - loss: 0.2623 - sparse_categorical_accuracy: 0.9202 - val_loss: 0.1836 - val_sparse_ca
tegorical_accuracy: 0.9488 - 2s/epoch - 3ms/step
Epoch 57/100
469/469 - 1s - loss: 0.2673 - sparse_categorical_accuracy: 0.9177 - val_loss: 0.1886 - val_sparse_ca
tegorical_accuracy: 0.9477 - 1s/epoch - 3ms/step
Epoch 58/100
469/469 - 1s - loss: 0.2643 - sparse_categorical_accuracy: 0.9183 - val_loss: 0.1899 - val_sparse_ca
tegorical_accuracy: 0.9450 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.2652 - sparse_categorical_accuracy: 0.9186 - val_loss: 0.2066 - val_sparse_ca
tegorical_accuracy: 0.9410 - 1s/epoch - 3ms/step
Epoch 60/100
469/469 - 2s - loss: 0.2656 - sparse_categorical_accuracy: 0.9186 - val_loss: 0.1976 - val_sparse_ca
tegorical_accuracy: 0.9450 - 2s/epoch - 5ms/step
Epoch 61/100
469/469 - 2s - loss: 0.2654 - sparse_categorical_accuracy: 0.9196 - val_loss: 0.1858 - val_sparse_ca
tegorical_accuracy: 0.9468 - 2s/epoch - 5ms/step
Epoch 62/100
469/469 - 2s - loss: 0.2604 - sparse_categorical_accuracy: 0.9198 - val_loss: 0.1902 - val_sparse_ca
tegorical_accuracy: 0.9468 - 2s/epoch - 5ms/step
Epoch 63/100
469/469 - 2s - loss: 0.2615 - sparse_categorical_accuracy: 0.9198 - val_loss: 0.2047 - val_sparse_ca
tegorical_accuracy: 0.9420 - 2s/epoch - 5ms/step
Epoch 64/100
469/469 - 2s - loss: 0.2615 - sparse_categorical_accuracy: 0.9201 - val_loss: 0.2019 - val_sparse_ca
tegorical_accuracy: 0.9436 - 2s/epoch - 5ms/step
Epoch 65/100
469/469 - 2s - loss: 0.2616 - sparse_categorical_accuracy: 0.9194 - val_loss: 0.1896 - val_sparse_ca
tegorical_accuracy: 0.9463 - 2s/epoch - 5ms/step
Epoch 66/100
469/469 - 2s - loss: 0.2533 - sparse_categorical_accuracy: 0.9223 - val_loss: 0.1891 - val_sparse_ca
tegorical_accuracy: 0.9484 - 2s/epoch - 4ms/step
Epoch 67/100
469/469 - 2s - loss: 0.2554 - sparse_categorical_accuracy: 0.9216 - val_loss: 0.1960 - val_sparse_ca
tegorical_accuracy: 0.9417 - 2s/epoch - 4ms/step
Epoch 68/100
469/469 - 2s - loss: 0.2607 - sparse_categorical_accuracy: 0.9201 - val_loss: 0.2080 - val_sparse_ca
tegorical_accuracy: 0.9444 - 2s/epoch - 5ms/step
Epoch 69/100
469/469 - 2s - loss: 0.2610 - sparse_categorical_accuracy: 0.9193 - val_loss: 0.1896 - val_sparse_ca
tegorical_accuracy: 0.9468 - 2s/epoch - 4ms/step
Epoch 70/100
```

```
469/469 - 2s - loss: 0.2562 - sparse_categorical_accuracy: 0.9206 - val_loss: 0.1926 - val_sparse_ca
tegorical_accuracy: 0.9466 - 2s/epoch - 4ms/step
Epoch 71/100
469/469 - 2s - loss: 0.2609 - sparse_categorical_accuracy: 0.9202 - val_loss: 0.2102 - val_sparse_ca
tegorical_accuracy: 0.9416 - 2s/epoch - 4ms/step
Epoch 72/100
469/469 - 2s - loss: 0.2590 - sparse_categorical_accuracy: 0.9193 - val_loss: 0.2014 - val_sparse_ca
tegorical_accuracy: 0.9448 - 2s/epoch - 4ms/step
Epoch 73/100
469/469 - 2s - loss: 0.2541 - sparse_categorical_accuracy: 0.9215 - val_loss: 0.1991 - val_sparse_ca
tegorical_accuracy: 0.9439 - 2s/epoch - 4ms/step
Epoch 74/100
469/469 - 1s - loss: 0.2607 - sparse_categorical_accuracy: 0.9198 - val_loss: 0.1998 - val_sparse_ca
tegorical_accuracy: 0.9437 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.2643 - sparse_categorical_accuracy: 0.9190 - val_loss: 0.2047 - val_sparse_ca
tegorical_accuracy: 0.9427 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.2580 - sparse_categorical_accuracy: 0.9203 - val_loss: 0.1960 - val_sparse_ca
tegorical_accuracy: 0.9423 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.2544 - sparse_categorical_accuracy: 0.9221 - val_loss: 0.1882 - val_sparse_ca
tegorical_accuracy: 0.9487 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.2559 - sparse_categorical_accuracy: 0.9219 - val_loss: 0.2055 - val_sparse_ca
tegorical_accuracy: 0.9432 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.2579 - sparse_categorical_accuracy: 0.9210 - val_loss: 0.1989 - val_sparse_ca
tegorical_accuracy: 0.9445 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.2563 - sparse_categorical_accuracy: 0.9215 - val_loss: 0.1951 - val_sparse_ca
tegorical_accuracy: 0.9450 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.2530 - sparse_categorical_accuracy: 0.9216 - val_loss: 0.1869 - val_sparse_ca
tegorical_accuracy: 0.9475 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.2562 - sparse_categorical_accuracy: 0.9218 - val_loss: 0.1941 - val_sparse_ca
tegorical_accuracy: 0.9478 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.2534 - sparse_categorical_accuracy: 0.9230 - val_loss: 0.2007 - val_sparse_ca
tegorical_accuracy: 0.9458 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.2533 - sparse_categorical_accuracy: 0.9226 - val_loss: 0.1905 - val_sparse_ca
tegorical_accuracy: 0.9465 - 1s/epoch - 3ms/step
Epoch 85/100
469/469 - 1s - loss: 0.2532 - sparse_categorical_accuracy: 0.9232 - val_loss: 0.1961 - val_sparse_ca
tegorical_accuracy: 0.9444 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.2559 - sparse_categorical_accuracy: 0.9209 - val_loss: 0.2052 - val_sparse_ca
tegorical_accuracy: 0.9434 - 1s/epoch - 3ms/step
Epoch 87/100
469/469 - 1s - loss: 0.2498 - sparse_categorical_accuracy: 0.9232 - val_loss: 0.1951 - val_sparse_ca
tegorical_accuracy: 0.9443 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.2556 - sparse_categorical_accuracy: 0.9208 - val_loss: 0.1893 - val_sparse_ca
tegorical_accuracy: 0.9448 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.2560 - sparse_categorical_accuracy: 0.9211 - val_loss: 0.1929 - val_sparse_ca
tegorical_accuracy: 0.9435 - 1s/epoch - 3ms/step
Epoch 90/100
469/469 - 1s - loss: 0.2545 - sparse_categorical_accuracy: 0.9218 - val_loss: 0.2026 - val_sparse_ca
tegorical_accuracy: 0.9435 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.2507 - sparse_categorical_accuracy: 0.9223 - val_loss: 0.2146 - val_sparse_ca
tegorical_accuracy: 0.9451 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.2532 - sparse_categorical_accuracy: 0.9218 - val_loss: 0.2009 - val_sparse_ca
tegorical_accuracy: 0.9429 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.2521 - sparse_categorical_accuracy: 0.9228 - val_loss: 0.1940 - val_sparse_ca
tegorical_accuracy: 0.9441 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.2556 - sparse_categorical_accuracy: 0.9219 - val_loss: 0.1907 - val_sparse_ca
tegorical_accuracy: 0.9462 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.2500 - sparse_categorical_accuracy: 0.9241 - val_loss: 0.1879 - val_sparse_ca
tegorical_accuracy: 0.9481 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.2546 - sparse_categorical_accuracy: 0.9215 - val_loss: 0.1908 - val_sparse_ca
tegorical_accuracy: 0.9453 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.2510 - sparse_categorical_accuracy: 0.9233 - val_loss: 0.1913 - val_sparse_ca
tegorical_accuracy: 0.9450 - 1s/epoch - 3ms/step
```

```
Epoch 98/100
469/469 - 1s - loss: 0.2537 - sparse_categorical_accuracy: 0.9226 - val_loss: 0.2004 - val_sparse_ca
tegorical_accuracy: 0.9428 - 1s/epoch - 3ms/step
Epoch 99/100
469/469 - 1s - loss: 0.2507 - sparse_categorical_accuracy: 0.9228 - val_loss: 0.1936 - val_sparse_ca
tegorical_accuracy: 0.9489 - 1s/epoch - 3ms/step
Epoch 100/100
469/469 - 1s - loss: 0.2519 - sparse_categorical_accuracy: 0.9225 - val_loss: 0.1892 - val_sparse_ca
tegorical_accuracy: 0.9454 - 1s/epoch - 2ms/step
```

In [86]:

```
model9.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.1892 - sparse_categorical_accuracy: 0.9454 - 407ms/epoch - 1ms/step
```

Out[86]:

```
[0.18921124935150146, 0.9453999996185303]
```

Model10: Dropout with parameter 0.3

In [87]:

```
model10 = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dropout(0.3),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
model10.compile(optimizer=tf.keras.optimizers.Adam(0.001),
                loss='sparse_categorical_crossentropy',
                metrics=['sparse_categorical_accuracy']
                )
history10 = model10.fit(x_train, y_train,
                        batch_size=128,
                        epochs=100,
                        validation_data=(x_test, y_test),
                        verbose=2
                        )
```

```
Epoch 1/100
469/469 - 2s - loss: 0.7978 - sparse_categorical_accuracy: 0.7264 - val_loss: 0.3104 - val_sparse_ca
tegorical_accuracy: 0.9121 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 0.5553 - sparse_categorical_accuracy: 0.8130 - val_loss: 0.2705 - val_sparse_ca
tegorical_accuracy: 0.9238 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 0.4970 - sparse_categorical_accuracy: 0.8365 - val_loss: 0.2440 - val_sparse_ca
tegorical_accuracy: 0.9307 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 0.4756 - sparse_categorical_accuracy: 0.8449 - val_loss: 0.2500 - val_sparse_ca
tegorical_accuracy: 0.9238 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 0.4551 - sparse_categorical_accuracy: 0.8514 - val_loss: 0.2489 - val_sparse_ca
tegorical_accuracy: 0.9271 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 0.4481 - sparse_categorical_accuracy: 0.8541 - val_loss: 0.2449 - val_sparse_ca
tegorical_accuracy: 0.9234 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.4486 - sparse_categorical_accuracy: 0.8540 - val_loss: 0.2296 - val_sparse_ca
tegorical_accuracy: 0.9318 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 0.4385 - sparse_categorical_accuracy: 0.8559 - val_loss: 0.2536 - val_sparse_ca
tegorical_accuracy: 0.9260 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 0.4321 - sparse_categorical_accuracy: 0.8569 - val_loss: 0.2433 - val_sparse_ca
tegorical_accuracy: 0.9303 - 1s/epoch - 3ms/step
Epoch 10/100
469/469 - 1s - loss: 0.4313 - sparse_categorical_accuracy: 0.8598 - val_loss: 0.2253 - val_sparse_ca
tegorical_accuracy: 0.9328 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 0.4295 - sparse_categorical_accuracy: 0.8605 - val_loss: 0.2280 - val_sparse_ca
tegorical_accuracy: 0.9317 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.4270 - sparse_categorical_accuracy: 0.8597 - val_loss: 0.2341 - val_sparse_ca
tegorical_accuracy: 0.9317 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.4205 - sparse_categorical_accuracy: 0.8640 - val_loss: 0.2487 - val_sparse_ca
tegorical_accuracy: 0.9277 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.4178 - sparse_categorical_accuracy: 0.8648 - val_loss: 0.2278 - val_sparse_ca
```

```
tegorical_accuracy: 0.9352 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 0.4130 - sparse_categorical_accuracy: 0.8655 - val_loss: 0.2361 - val_sparse_ca
tegorical_accuracy: 0.9338 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.4149 - sparse_categorical_accuracy: 0.8641 - val_loss: 0.2395 - val_sparse_ca
tegorical_accuracy: 0.9308 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.4129 - sparse_categorical_accuracy: 0.8660 - val_loss: 0.2225 - val_sparse_ca
tegorical_accuracy: 0.9354 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 0.4097 - sparse_categorical_accuracy: 0.8663 - val_loss: 0.2306 - val_sparse_ca
tegorical_accuracy: 0.9330 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.4062 - sparse_categorical_accuracy: 0.8682 - val_loss: 0.2467 - val_sparse_ca
tegorical_accuracy: 0.9286 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.4009 - sparse_categorical_accuracy: 0.8700 - val_loss: 0.2602 - val_sparse_ca
tegorical_accuracy: 0.9250 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.3954 - sparse_categorical_accuracy: 0.8708 - val_loss: 0.2277 - val_sparse_ca
tegorical_accuracy: 0.9356 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.4086 - sparse_categorical_accuracy: 0.8684 - val_loss: 0.2576 - val_sparse_ca
tegorical_accuracy: 0.9251 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.3974 - sparse_categorical_accuracy: 0.8726 - val_loss: 0.2395 - val_sparse_ca
tegorical_accuracy: 0.9307 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.3932 - sparse_categorical_accuracy: 0.8730 - val_loss: 0.2236 - val_sparse_ca
tegorical_accuracy: 0.9373 - 1s/epoch - 3ms/step
Epoch 25/100
469/469 - 1s - loss: 0.3936 - sparse_categorical_accuracy: 0.8723 - val_loss: 0.2356 - val_sparse_ca
tegorical_accuracy: 0.9313 - 1s/epoch - 3ms/step
Epoch 26/100
469/469 - 1s - loss: 0.3905 - sparse_categorical_accuracy: 0.8734 - val_loss: 0.2340 - val_sparse_ca
tegorical_accuracy: 0.9358 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.3862 - sparse_categorical_accuracy: 0.8741 - val_loss: 0.2526 - val_sparse_ca
tegorical_accuracy: 0.9263 - 1s/epoch - 3ms/step
Epoch 28/100
469/469 - 1s - loss: 0.3901 - sparse_categorical_accuracy: 0.8728 - val_loss: 0.2190 - val_sparse_ca
tegorical_accuracy: 0.9393 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.3878 - sparse_categorical_accuracy: 0.8759 - val_loss: 0.2323 - val_sparse_ca
tegorical_accuracy: 0.9318 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.3887 - sparse_categorical_accuracy: 0.8757 - val_loss: 0.2419 - val_sparse_ca
tegorical_accuracy: 0.9311 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.3888 - sparse_categorical_accuracy: 0.8750 - val_loss: 0.2373 - val_sparse_ca
tegorical_accuracy: 0.9337 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.3927 - sparse_categorical_accuracy: 0.8740 - val_loss: 0.2451 - val_sparse_ca
tegorical_accuracy: 0.9305 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.3840 - sparse_categorical_accuracy: 0.8766 - val_loss: 0.2368 - val_sparse_ca
tegorical_accuracy: 0.9309 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.3867 - sparse_categorical_accuracy: 0.8739 - val_loss: 0.2199 - val_sparse_ca
tegorical_accuracy: 0.9404 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.3855 - sparse_categorical_accuracy: 0.8752 - val_loss: 0.2347 - val_sparse_ca
tegorical_accuracy: 0.9325 - 1s/epoch - 3ms/step
Epoch 36/100
469/469 - 1s - loss: 0.3806 - sparse_categorical_accuracy: 0.8767 - val_loss: 0.2261 - val_sparse_ca
tegorical_accuracy: 0.9355 - 1s/epoch - 3ms/step
Epoch 37/100
469/469 - 1s - loss: 0.3818 - sparse_categorical_accuracy: 0.8768 - val_loss: 0.2460 - val_sparse_ca
tegorical_accuracy: 0.9331 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.3764 - sparse_categorical_accuracy: 0.8786 - val_loss: 0.2195 - val_sparse_ca
tegorical_accuracy: 0.9347 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.3801 - sparse_categorical_accuracy: 0.8785 - val_loss: 0.2178 - val_sparse_ca
tegorical_accuracy: 0.9396 - 1s/epoch - 3ms/step
Epoch 40/100
469/469 - 1s - loss: 0.3845 - sparse_categorical_accuracy: 0.8758 - val_loss: 0.2280 - val_sparse_ca
tegorical_accuracy: 0.9360 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.3842 - sparse_categorical_accuracy: 0.8771 - val_loss: 0.2283 - val_sparse_ca
tegorical_accuracy: 0.9381 - 1s/epoch - 2ms/step
Epoch 42/100
```

```
469/469 - 1s - loss: 0.3780 - sparse_categorical_accuracy: 0.8767 - val_loss: 0.2297 - val_sparse_ca
tegorical_accuracy: 0.9365 - 1s/epoch - 3ms/step
Epoch 43/100
469/469 - 1s - loss: 0.3763 - sparse_categorical_accuracy: 0.8784 - val_loss: 0.2358 - val_sparse_ca
tegorical_accuracy: 0.9369 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.3803 - sparse_categorical_accuracy: 0.8779 - val_loss: 0.2331 - val_sparse_ca
tegorical_accuracy: 0.9320 - 1s/epoch - 3ms/step
Epoch 45/100
469/469 - 1s - loss: 0.3742 - sparse_categorical_accuracy: 0.8810 - val_loss: 0.2195 - val_sparse_ca
tegorical_accuracy: 0.9381 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.3771 - sparse_categorical_accuracy: 0.8792 - val_loss: 0.2292 - val_sparse_ca
tegorical_accuracy: 0.9353 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.3806 - sparse_categorical_accuracy: 0.8778 - val_loss: 0.2217 - val_sparse_ca
tegorical_accuracy: 0.9389 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.3772 - sparse_categorical_accuracy: 0.8781 - val_loss: 0.2271 - val_sparse_ca
tegorical_accuracy: 0.9360 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.3757 - sparse_categorical_accuracy: 0.8793 - val_loss: 0.2327 - val_sparse_ca
tegorical_accuracy: 0.9345 - 1s/epoch - 3ms/step
Epoch 50/100
469/469 - 1s - loss: 0.3755 - sparse_categorical_accuracy: 0.8802 - val_loss: 0.2522 - val_sparse_ca
tegorical_accuracy: 0.9283 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.3780 - sparse_categorical_accuracy: 0.8786 - val_loss: 0.2185 - val_sparse_ca
tegorical_accuracy: 0.9398 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.3742 - sparse_categorical_accuracy: 0.8810 - val_loss: 0.2289 - val_sparse_ca
tegorical_accuracy: 0.9350 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.3723 - sparse_categorical_accuracy: 0.8808 - val_loss: 0.2357 - val_sparse_ca
tegorical_accuracy: 0.9343 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.3749 - sparse_categorical_accuracy: 0.8809 - val_loss: 0.2495 - val_sparse_ca
tegorical_accuracy: 0.9280 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.3750 - sparse_categorical_accuracy: 0.8812 - val_loss: 0.2347 - val_sparse_ca
tegorical_accuracy: 0.9332 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 2s - loss: 0.3705 - sparse_categorical_accuracy: 0.8804 - val_loss: 0.2286 - val_sparse_ca
tegorical_accuracy: 0.9355 - 2s/epoch - 3ms/step
Epoch 57/100
469/469 - 1s - loss: 0.3739 - sparse_categorical_accuracy: 0.8811 - val_loss: 0.2258 - val_sparse_ca
tegorical_accuracy: 0.9358 - 1s/epoch - 3ms/step
Epoch 58/100
469/469 - 1s - loss: 0.3709 - sparse_categorical_accuracy: 0.8817 - val_loss: 0.2295 - val_sparse_ca
tegorical_accuracy: 0.9332 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.3696 - sparse_categorical_accuracy: 0.8821 - val_loss: 0.2341 - val_sparse_ca
tegorical_accuracy: 0.9372 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.3701 - sparse_categorical_accuracy: 0.8819 - val_loss: 0.2315 - val_sparse_ca
tegorical_accuracy: 0.9366 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.3653 - sparse_categorical_accuracy: 0.8833 - val_loss: 0.2355 - val_sparse_ca
tegorical_accuracy: 0.9320 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.3711 - sparse_categorical_accuracy: 0.8823 - val_loss: 0.2587 - val_sparse_ca
tegorical_accuracy: 0.9262 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 0.3680 - sparse_categorical_accuracy: 0.8819 - val_loss: 0.2277 - val_sparse_ca
tegorical_accuracy: 0.9376 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.3722 - sparse_categorical_accuracy: 0.8823 - val_loss: 0.2257 - val_sparse_ca
tegorical_accuracy: 0.9370 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.3706 - sparse_categorical_accuracy: 0.8805 - val_loss: 0.2266 - val_sparse_ca
tegorical_accuracy: 0.9367 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.3693 - sparse_categorical_accuracy: 0.8827 - val_loss: 0.2456 - val_sparse_ca
tegorical_accuracy: 0.9293 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.3668 - sparse_categorical_accuracy: 0.8830 - val_loss: 0.2304 - val_sparse_ca
tegorical_accuracy: 0.9352 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 0.3681 - sparse_categorical_accuracy: 0.8825 - val_loss: 0.2206 - val_sparse_ca
tegorical_accuracy: 0.9361 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 0.3645 - sparse_categorical_accuracy: 0.8831 - val_loss: 0.2360 - val_sparse_ca
tegorical_accuracy: 0.9304 - 1s/epoch - 2ms/step
```

```
Epoch 70/100
469/469 - 1s - loss: 0.3647 - sparse_categorical_accuracy: 0.8838 - val_loss: 0.2315 - val_sparse_ca
tegorical_accuracy: 0.9337 - 1s/epoch - 3ms/step
Epoch 71/100
469/469 - 1s - loss: 0.3600 - sparse_categorical_accuracy: 0.8845 - val_loss: 0.2284 - val_sparse_ca
tegorical_accuracy: 0.9366 - 1s/epoch - 3ms/step
Epoch 72/100
469/469 - 1s - loss: 0.3627 - sparse_categorical_accuracy: 0.8847 - val_loss: 0.2191 - val_sparse_ca
tegorical_accuracy: 0.9401 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.3656 - sparse_categorical_accuracy: 0.8829 - val_loss: 0.2358 - val_sparse_ca
tegorical_accuracy: 0.9343 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.3643 - sparse_categorical_accuracy: 0.8842 - val_loss: 0.2306 - val_sparse_ca
tegorical_accuracy: 0.9343 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.3638 - sparse_categorical_accuracy: 0.8842 - val_loss: 0.2197 - val_sparse_ca
tegorical_accuracy: 0.9385 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.3604 - sparse_categorical_accuracy: 0.8863 - val_loss: 0.2418 - val_sparse_ca
tegorical_accuracy: 0.9335 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.3631 - sparse_categorical_accuracy: 0.8849 - val_loss: 0.2333 - val_sparse_ca
tegorical_accuracy: 0.9327 - 1s/epoch - 3ms/step
Epoch 78/100
469/469 - 1s - loss: 0.3609 - sparse_categorical_accuracy: 0.8859 - val_loss: 0.2221 - val_sparse_ca
tegorical_accuracy: 0.9370 - 1s/epoch - 3ms/step
Epoch 79/100
469/469 - 1s - loss: 0.3615 - sparse_categorical_accuracy: 0.8859 - val_loss: 0.2226 - val_sparse_ca
tegorical_accuracy: 0.9383 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.3611 - sparse_categorical_accuracy: 0.8866 - val_loss: 0.2286 - val_sparse_ca
tegorical_accuracy: 0.9329 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.3656 - sparse_categorical_accuracy: 0.8831 - val_loss: 0.2491 - val_sparse_ca
tegorical_accuracy: 0.9295 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.3600 - sparse_categorical_accuracy: 0.8849 - val_loss: 0.2363 - val_sparse_ca
tegorical_accuracy: 0.9309 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.3619 - sparse_categorical_accuracy: 0.8846 - val_loss: 0.2255 - val_sparse_ca
tegorical_accuracy: 0.9369 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.3622 - sparse_categorical_accuracy: 0.8859 - val_loss: 0.2207 - val_sparse_ca
tegorical_accuracy: 0.9356 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.3601 - sparse_categorical_accuracy: 0.8858 - val_loss: 0.2378 - val_sparse_ca
tegorical_accuracy: 0.9317 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.3605 - sparse_categorical_accuracy: 0.8860 - val_loss: 0.2262 - val_sparse_ca
tegorical_accuracy: 0.9377 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.3622 - sparse_categorical_accuracy: 0.8850 - val_loss: 0.2290 - val_sparse_ca
tegorical_accuracy: 0.9350 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.3650 - sparse_categorical_accuracy: 0.8844 - val_loss: 0.2254 - val_sparse_ca
tegorical_accuracy: 0.9334 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.3608 - sparse_categorical_accuracy: 0.8865 - val_loss: 0.2412 - val_sparse_ca
tegorical_accuracy: 0.9273 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.3600 - sparse_categorical_accuracy: 0.8862 - val_loss: 0.2286 - val_sparse_ca
tegorical_accuracy: 0.9341 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.3616 - sparse_categorical_accuracy: 0.8867 - val_loss: 0.2232 - val_sparse_ca
tegorical_accuracy: 0.9379 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.3568 - sparse_categorical_accuracy: 0.8871 - val_loss: 0.2291 - val_sparse_ca
tegorical_accuracy: 0.9343 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.3593 - sparse_categorical_accuracy: 0.8871 - val_loss: 0.2406 - val_sparse_ca
tegorical_accuracy: 0.9328 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.3571 - sparse_categorical_accuracy: 0.8875 - val_loss: 0.2255 - val_sparse_ca
tegorical_accuracy: 0.9345 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.3553 - sparse_categorical_accuracy: 0.8894 - val_loss: 0.2333 - val_sparse_ca
tegorical_accuracy: 0.9331 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.3577 - sparse_categorical_accuracy: 0.8859 - val_loss: 0.2398 - val_sparse_ca
tegorical_accuracy: 0.9302 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.3547 - sparse_categorical_accuracy: 0.8867 - val_loss: 0.2325 - val_sparse_ca
```

```
tegorical_accuracy: 0.9330 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.3523 - sparse_categorical_accuracy: 0.8884 - val_loss: 0.2346 - val_sparse_ca
tegorical_accuracy: 0.9319 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.3615 - sparse_categorical_accuracy: 0.8848 - val_loss: 0.2312 - val_sparse_ca
tegorical_accuracy: 0.9300 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.3560 - sparse_categorical_accuracy: 0.8870 - val_loss: 0.2645 - val_sparse_ca
tegorical_accuracy: 0.9218 - 1s/epoch - 2ms/step
```

In [88]:

```
model10.evaluate(x_test, y_test, verbose=2)
```

```
313/313 - 0s - loss: 0.2645 - sparse_categorical_accuracy: 0.9218 - 401ms/epoch - 1ms/step
```

Out[88]:

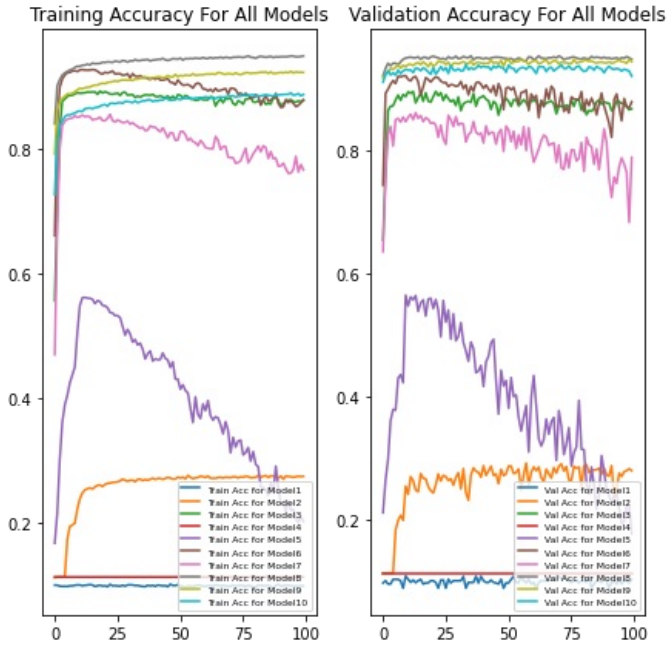
```
[0.26446375250816345, 0.9218000173568726]
```

Plots

Plots for different regularization setups

In [92]:

```
training_accuracy1 = history1.history['sparse_categorical_accuracy']
validation_accuracy1 = history1.history['val_sparse_categorical_accuracy']

training_accuracy2 = history2.history['sparse_categorical_accuracy']
validation_accuracy2 = history2.history['val_sparse_categorical_accuracy']

training_accuracy3 = history3.history['sparse_categorical_accuracy']
validation_accuracy3 = history3.history['val_sparse_categorical_accuracy']

training_accuracy4 = history4.history['sparse_categorical_accuracy']
validation_accuracy4 = history4.history['val_sparse_categorical_accuracy']

training_accuracy5 = history5.history['sparse_categorical_accuracy']
validation_accuracy5 = history5.history['val_sparse_categorical_accuracy']

training_accuracy6 = history6.history['sparse_categorical_accuracy']
validation_accuracy6 = history6.history['val_sparse_categorical_accuracy']

training_accuracy7 = history7.history['sparse_categorical_accuracy']
validation_accuracy7 = history7.history['val_sparse_categorical_accuracy']

training_accuracy8 = history8.history['sparse_categorical_accuracy']
validation_accuracy8 = history8.history['val_sparse_categorical_accuracy']

training_accuracy9 = history9.history['sparse_categorical_accuracy']
validation_accuracy9 = history9.history['val_sparse_categorical_accuracy']

training_accuracy10 = history10.history['sparse_categorical_accuracy']
validation_accuracy10 = history10.history['val_sparse_categorical_accuracy']

epochs_range=range(100)

plt.figure(figsize=(7, 7))
plt.subplot(1, 2, 1)
plt.plot(epochs_range, training_accuracy1, label='Train Acc for Model1')
plt.plot(epochs_range, training_accuracy2, label='Train Acc for Model2')
plt.plot(epochs_range, training_accuracy3, label='Train Acc for Model3')
plt.plot(epochs_range, training_accuracy4, label='Train Acc for Model4')
plt.plot(epochs_range, training_accuracy5, label='Train Acc for Model5')
plt.plot(epochs_range, training_accuracy6, label='Train Acc for Model6')
plt.plot(epochs_range, training_accuracy7, label='Train Acc for Model7')
plt.plot(epochs_range, training_accuracy8, label='Train Acc for Model8')
plt.plot(epochs_range, training_accuracy9, label='Train Acc for Model9')
plt.plot(epochs_range, training_accuracy10, label='Train Acc for Model10')
plt.legend(loc='lower right', prop={'size': 6})
plt.title('Training Accuracy For All Models')

plt.subplot(1, 2, 2)
plt.plot(epochs_range, validation_accuracy1, label='Val Acc for Model1')
plt.plot(epochs_range, validation_accuracy2, label='Val Acc for Model2')
plt.plot(epochs_range, validation_accuracy3, label='Val Acc for Model3')
plt.plot(epochs_range, validation_accuracy4, label='Val Acc for Model4')
plt.plot(epochs_range, validation_accuracy5, label='Val Acc for Model5')
plt.plot(epochs_range, validation_accuracy6, label='Val Acc for Model6')
plt.plot(epochs_range, validation_accuracy7, label='Val Acc for Model7')
plt.plot(epochs_range, validation_accuracy8, label='Val Acc for Model8')
plt.plot(epochs_range, validation_accuracy9, label='Val Acc for Model9')
plt.plot(epochs_range, validation_accuracy10, label='Val Acc for Model10')
plt.legend(loc='lower right', prop={'size': 6})
plt.title('Validation Accuracy For All Models')
plt.show()
```



Analysis of plot

Both model1 and model7 has extremely low traning and validation accuracy through out the training process. That's because penalty parameter is too high in both cases so that loss function is dominated solely by sum of (absolute values of)weights, where model can simply make its weights become all zero to achieve the smallest loss ignoring the classification question.</br> Therefore, regularization parameters shouldn't be too large in which case training becomes travial. We can see that model3 and model6 have good performances, I believe it means their choices of parameters are good: it's not too large to override the training, nor too small such that the regulrization penalty has no impact on training at all.</br> One thing to notice is that L1 and/or L2 regularization can indeed prevent overfitting, since training accuracies don't keep increasing with epoches in model1 ~ model8.</br> All dropout setups can produce excellent performances but it can't prevent overfitting, as validation accuracies don't keep increasing with epoches.

Best model in MNIST hand-written digits dataset

Both model3 and model6 has good performances, since model6 outperforms model3 in most of the epoches, I would choose model3 to be the best L1 and/or L2 regularization model. But taking performances of dropout into account, obviously model8 would be the best model.

Diving into fashion-MNIST dataset

Model1: Regularization penalty is L1 with parameter 10^{-1}

In [62]:

```
from keras.datasets import fashion_mnist
(trainX_f, trainy_f), (testX_f, testy_f) = 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 [101]:

```
model1_f = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform',
  activity_regularizer = tf.keras.regularizers.L1(0.1)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
```

In [102]:

```
# Fix the learning rate to be 0.001
model1_f.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
```

In [103]:

```
history1_f = model1_f.fit(trainX_f, trainy_f,
                      batch_size=128,
                      epochs=100,
                      validation_data=(testX_f, testy_f),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 2s - loss: 2.5451 - sparse_categorical_accuracy: 0.0960 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 2s/epoch - 5ms/step
Epoch 2/100
469/469 - 2s - loss: 2.3028 - sparse_categorical_accuracy: 0.0997 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 2s/epoch - 3ms/step
Epoch 3/100
469/469 - 1s - loss: 2.3028 - sparse_categorical_accuracy: 0.0983 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 4/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0992 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 5/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0978 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 6/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0992 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 7/100
469/469 - 2s - loss: 2.3027 - sparse_categorical_accuracy: 0.0977 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 2s/epoch - 3ms/step
Epoch 8/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.1006 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 9/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0974 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 10/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0966 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 11/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0987 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0995 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 13/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0978 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 14/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0978 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 15/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.1003 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 16/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0994 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 17/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0964 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 18/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0988 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 19/100
```

```
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0980 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0967 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0988 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0957 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0973 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0966 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 25/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0982 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 26/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0982 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 27/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0963 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0989 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0976 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0966 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0974 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0988 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0983 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0976 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0980 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 36/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0973 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0959 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0970 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 39/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0978 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0981 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0993 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 42/100
469/469 - 2s - loss: 2.3027 - sparse_categorical_accuracy: 0.0950 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 2s/epoch - 4ms/step
Epoch 43/100
469/469 - 2s - loss: 2.3027 - sparse_categorical_accuracy: 0.0978 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 2s/epoch - 4ms/step
Epoch 44/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0986 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 2s - loss: 2.3027 - sparse_categorical_accuracy: 0.0974 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 2s/epoch - 3ms/step
Epoch 46/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0972 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
```

```
Epoch 47/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0988 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0974 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0981 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0986 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0975 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0989 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0967 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0981 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0972 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0992 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0995 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0981 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 59/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0974 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0975 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0981 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0968 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0980 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0975 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0981 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0988 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0967 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0983 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0992 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 3ms/step
Epoch 72/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0984 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0958 - val_loss: 2.3027 - val_sparse_ca
```

```
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0990 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0983 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0984 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0982 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0981 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0993 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0987 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0997 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0990 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0970 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0983 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0965 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0972 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0984 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0994 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0973 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0988 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0992 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 94/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0982 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 95/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0975 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0981 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0990 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0963 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0983 - val_loss: 2.3027 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
```

In [104]:

```
model1_f.evaluate(testX_f, testy_f, verbose=2)
```

```
313/313 - 0s - loss: 2.3027 - sparse_categorical_accuracy: 0.1001 - 392ms/epoch - 1ms/step
```

Out[104]:

```
[2.302670478820801, 0.10010000318288803]
```

Model2: Regularization penalty is L1 with parameter 10^{-2}

In [105]:

```
model2_f = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform',
  activity_regularizer = tf.keras.regularizers.L1(0.01)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
# Fix the learning rate to be 0.001
model2_f.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
history2_f = model2_f.fit(trainX_f, trainy_f,
                     batch_size=128,
                     epochs=100,
                     validation_data=(testX_f, testy_f),
                     verbose=2
                     )
```

```
Epoch 1/100
469/469 - 2s - loss: 2.3280 - sparse_categorical_accuracy: 0.0993 - val_loss: 2.3026 - val_sparse_ca
tegorical_accuracy: 0.1000 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0991 - val_loss: 2.3026 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0979 - val_loss: 2.3026 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 2.2846 - sparse_categorical_accuracy: 0.1112 - val_loss: 2.2954 - val_sparse_ca
tegorical_accuracy: 0.1054 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 2.1874 - sparse_categorical_accuracy: 0.1672 - val_loss: 2.1623 - val_sparse_ca
tegorical_accuracy: 0.1742 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 2.0884 - sparse_categorical_accuracy: 0.1904 - val_loss: 2.0273 - val_sparse_ca
tegorical_accuracy: 0.1994 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 2.0142 - sparse_categorical_accuracy: 0.1992 - val_loss: 1.9504 - val_sparse_ca
tegorical_accuracy: 0.1996 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 1.9699 - sparse_categorical_accuracy: 0.1906 - val_loss: 1.9255 - val_sparse_ca
tegorical_accuracy: 0.2035 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 1.9312 - sparse_categorical_accuracy: 0.1855 - val_loss: 1.8633 - val_sparse_ca
tegorical_accuracy: 0.1998 - 1s/epoch - 3ms/step
Epoch 10/100
469/469 - 1s - loss: 1.8979 - sparse_categorical_accuracy: 0.2115 - val_loss: 1.9658 - val_sparse_ca
tegorical_accuracy: 0.2089 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 1.8599 - sparse_categorical_accuracy: 0.2339 - val_loss: 1.9407 - val_sparse_ca
tegorical_accuracy: 0.2020 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 1.8413 - sparse_categorical_accuracy: 0.2518 - val_loss: 1.8338 - val_sparse_ca
tegorical_accuracy: 0.2406 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 1.8183 - sparse_categorical_accuracy: 0.2539 - val_loss: 1.8793 - val_sparse_ca
tegorical_accuracy: 0.2475 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 1.8054 - sparse_categorical_accuracy: 0.2597 - val_loss: 1.8353 - val_sparse_ca
tegorical_accuracy: 0.2299 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 1.8032 - sparse_categorical_accuracy: 0.2596 - val_loss: 1.7632 - val_sparse_ca
tegorical_accuracy: 0.2590 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 1.7991 - sparse_categorical_accuracy: 0.2620 - val_loss: 1.7645 - val_sparse_ca
tegorical_accuracy: 0.2803 - 1s/epoch - 2ms/step
Epoch 17/100
```

```
469/469 - 1s - loss: 1.7982 - sparse_categorical_accuracy: 0.2612 - val_loss: 1.8286 - val_sparse_ca
tegorical_accuracy: 0.2736 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 1.7839 - sparse_categorical_accuracy: 0.2622 - val_loss: 1.8078 - val_sparse_ca
tegorical_accuracy: 0.2593 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 1.7785 - sparse_categorical_accuracy: 0.2654 - val_loss: 1.7272 - val_sparse_ca
tegorical_accuracy: 0.2678 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 1.8033 - sparse_categorical_accuracy: 0.2686 - val_loss: 1.7838 - val_sparse_ca
tegorical_accuracy: 0.2617 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 1.7611 - sparse_categorical_accuracy: 0.2705 - val_loss: 1.8000 - val_sparse_ca
tegorical_accuracy: 0.2715 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 1.7660 - sparse_categorical_accuracy: 0.2709 - val_loss: 1.7988 - val_sparse_ca
tegorical_accuracy: 0.2653 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 1.7739 - sparse_categorical_accuracy: 0.2711 - val_loss: 2.0456 - val_sparse_ca
tegorical_accuracy: 0.1787 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 1.7887 - sparse_categorical_accuracy: 0.2712 - val_loss: 1.9792 - val_sparse_ca
tegorical_accuracy: 0.2280 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 1.7814 - sparse_categorical_accuracy: 0.2680 - val_loss: 1.7401 - val_sparse_ca
tegorical_accuracy: 0.2790 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 1.7491 - sparse_categorical_accuracy: 0.2708 - val_loss: 1.8055 - val_sparse_ca
tegorical_accuracy: 0.2686 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 1.7822 - sparse_categorical_accuracy: 0.2718 - val_loss: 1.7618 - val_sparse_ca
tegorical_accuracy: 0.3062 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 1.7623 - sparse_categorical_accuracy: 0.2736 - val_loss: 1.7516 - val_sparse_ca
tegorical_accuracy: 0.3013 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 1.7462 - sparse_categorical_accuracy: 0.2769 - val_loss: 1.7465 - val_sparse_ca
tegorical_accuracy: 0.2565 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 1.7724 - sparse_categorical_accuracy: 0.2804 - val_loss: 1.7522 - val_sparse_ca
tegorical_accuracy: 0.2986 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 1.7640 - sparse_categorical_accuracy: 0.2783 - val_loss: 1.7196 - val_sparse_ca
tegorical_accuracy: 0.2677 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 1.7467 - sparse_categorical_accuracy: 0.2769 - val_loss: 1.7334 - val_sparse_ca
tegorical_accuracy: 0.2864 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 1.7478 - sparse_categorical_accuracy: 0.2791 - val_loss: 1.7416 - val_sparse_ca
tegorical_accuracy: 0.3007 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 1.7488 - sparse_categorical_accuracy: 0.2767 - val_loss: 1.7463 - val_sparse_ca
tegorical_accuracy: 0.2570 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 1.7505 - sparse_categorical_accuracy: 0.2807 - val_loss: 1.7861 - val_sparse_ca
tegorical_accuracy: 0.2956 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 1.7357 - sparse_categorical_accuracy: 0.2820 - val_loss: 1.7094 - val_sparse_ca
tegorical_accuracy: 0.2766 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 1.7460 - sparse_categorical_accuracy: 0.2794 - val_loss: 1.7361 - val_sparse_ca
tegorical_accuracy: 0.3010 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 1.7455 - sparse_categorical_accuracy: 0.2813 - val_loss: 1.7930 - val_sparse_ca
tegorical_accuracy: 0.3057 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 1.7396 - sparse_categorical_accuracy: 0.2831 - val_loss: 1.7153 - val_sparse_ca
tegorical_accuracy: 0.2975 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 1.7371 - sparse_categorical_accuracy: 0.2795 - val_loss: 1.7171 - val_sparse_ca
tegorical_accuracy: 0.2939 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 1.7631 - sparse_categorical_accuracy: 0.2796 - val_loss: 1.7996 - val_sparse_ca
tegorical_accuracy: 0.2665 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 1.7382 - sparse_categorical_accuracy: 0.2868 - val_loss: 1.7350 - val_sparse_ca
tegorical_accuracy: 0.2765 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 1.7341 - sparse_categorical_accuracy: 0.2845 - val_loss: 1.7740 - val_sparse_ca
tegorical_accuracy: 0.2778 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 1.7488 - sparse_categorical_accuracy: 0.2847 - val_loss: 1.7336 - val_sparse_ca
tegorical_accuracy: 0.3101 - 1s/epoch - 2ms/step
```

```
Epoch 45/100
469/469 - 1s - loss: 1.7386 - sparse_categorical_accuracy: 0.2882 - val_loss: 1.7130 - val_sparse_ca
tegorical_accuracy: 0.2883 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 1.7420 - sparse_categorical_accuracy: 0.2852 - val_loss: 1.7375 - val_sparse_ca
tegorical_accuracy: 0.3078 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 1.7375 - sparse_categorical_accuracy: 0.2887 - val_loss: 1.7417 - val_sparse_ca
tegorical_accuracy: 0.2593 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 1.7458 - sparse_categorical_accuracy: 0.2871 - val_loss: 1.7250 - val_sparse_ca
tegorical_accuracy: 0.3171 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 1.7433 - sparse_categorical_accuracy: 0.2860 - val_loss: 1.7068 - val_sparse_ca
tegorical_accuracy: 0.2862 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 1.7841 - sparse_categorical_accuracy: 0.2791 - val_loss: 1.7399 - val_sparse_ca
tegorical_accuracy: 0.2831 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 1.7623 - sparse_categorical_accuracy: 0.2830 - val_loss: 1.7235 - val_sparse_ca
tegorical_accuracy: 0.2814 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 1.7711 - sparse_categorical_accuracy: 0.2823 - val_loss: 1.7131 - val_sparse_ca
tegorical_accuracy: 0.2989 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 1.7424 - sparse_categorical_accuracy: 0.2904 - val_loss: 1.7125 - val_sparse_ca
tegorical_accuracy: 0.3053 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 1.7434 - sparse_categorical_accuracy: 0.2861 - val_loss: 1.7284 - val_sparse_ca
tegorical_accuracy: 0.2926 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 1.7514 - sparse_categorical_accuracy: 0.2830 - val_loss: 1.7718 - val_sparse_ca
tegorical_accuracy: 0.2813 - 1s/epoch - 3ms/step
Epoch 56/100
469/469 - 1s - loss: 1.7624 - sparse_categorical_accuracy: 0.2824 - val_loss: 1.7359 - val_sparse_ca
tegorical_accuracy: 0.2883 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 1.7556 - sparse_categorical_accuracy: 0.2827 - val_loss: 1.7313 - val_sparse_ca
tegorical_accuracy: 0.3197 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 1.7523 - sparse_categorical_accuracy: 0.2855 - val_loss: 1.7174 - val_sparse_ca
tegorical_accuracy: 0.2940 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 1.7546 - sparse_categorical_accuracy: 0.2836 - val_loss: 1.7619 - val_sparse_ca
tegorical_accuracy: 0.2664 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 1.7544 - sparse_categorical_accuracy: 0.2818 - val_loss: 1.7740 - val_sparse_ca
tegorical_accuracy: 0.2756 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 1.7577 - sparse_categorical_accuracy: 0.2819 - val_loss: 1.7329 - val_sparse_ca
tegorical_accuracy: 0.3065 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 1.7583 - sparse_categorical_accuracy: 0.2821 - val_loss: 1.7138 - val_sparse_ca
tegorical_accuracy: 0.3172 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 1.7595 - sparse_categorical_accuracy: 0.2866 - val_loss: 1.7296 - val_sparse_ca
tegorical_accuracy: 0.2768 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 1.7638 - sparse_categorical_accuracy: 0.2812 - val_loss: 1.7676 - val_sparse_ca
tegorical_accuracy: 0.2983 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 1.7545 - sparse_categorical_accuracy: 0.2894 - val_loss: 1.7533 - val_sparse_ca
tegorical_accuracy: 0.3204 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 1.7561 - sparse_categorical_accuracy: 0.2826 - val_loss: 1.7297 - val_sparse_ca
tegorical_accuracy: 0.2835 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 1.7539 - sparse_categorical_accuracy: 0.2878 - val_loss: 1.7659 - val_sparse_ca
tegorical_accuracy: 0.3173 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 1.7484 - sparse_categorical_accuracy: 0.2815 - val_loss: 1.7664 - val_sparse_ca
tegorical_accuracy: 0.2711 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 1.7660 - sparse_categorical_accuracy: 0.2773 - val_loss: 1.7241 - val_sparse_ca
tegorical_accuracy: 0.2652 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 1.7581 - sparse_categorical_accuracy: 0.2808 - val_loss: 1.7842 - val_sparse_ca
tegorical_accuracy: 0.2897 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 1.7549 - sparse_categorical_accuracy: 0.2800 - val_loss: 1.7399 - val_sparse_ca
tegorical_accuracy: 0.2794 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 1.7756 - sparse_categorical_accuracy: 0.2794 - val_loss: 1.8001 - val_sparse_ca
```

```
tegorical_accuracy: 0.2773 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 1.7558 - sparse_categorical_accuracy: 0.2786 - val_loss: 1.7203 - val_sparse_ca
tegorical_accuracy: 0.3026 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 1.7735 - sparse_categorical_accuracy: 0.2768 - val_loss: 1.7405 - val_sparse_ca
tegorical_accuracy: 0.3097 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 1.7611 - sparse_categorical_accuracy: 0.2851 - val_loss: 1.7752 - val_sparse_ca
tegorical_accuracy: 0.2849 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 1.7549 - sparse_categorical_accuracy: 0.2828 - val_loss: 1.7243 - val_sparse_ca
tegorical_accuracy: 0.2760 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 1.7503 - sparse_categorical_accuracy: 0.2834 - val_loss: 1.7306 - val_sparse_ca
tegorical_accuracy: 0.2724 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 1.7390 - sparse_categorical_accuracy: 0.2833 - val_loss: 1.7280 - val_sparse_ca
tegorical_accuracy: 0.2838 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 1.7659 - sparse_categorical_accuracy: 0.2786 - val_loss: 1.7801 - val_sparse_ca
tegorical_accuracy: 0.2957 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 1.7645 - sparse_categorical_accuracy: 0.2826 - val_loss: 1.8151 - val_sparse_ca
tegorical_accuracy: 0.2980 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 1.7912 - sparse_categorical_accuracy: 0.2740 - val_loss: 1.7783 - val_sparse_ca
tegorical_accuracy: 0.3020 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 1.7624 - sparse_categorical_accuracy: 0.2808 - val_loss: 1.7199 - val_sparse_ca
tegorical_accuracy: 0.2782 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 1.7782 - sparse_categorical_accuracy: 0.2773 - val_loss: 1.7268 - val_sparse_ca
tegorical_accuracy: 0.2814 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 1.7835 - sparse_categorical_accuracy: 0.2787 - val_loss: 1.7578 - val_sparse_ca
tegorical_accuracy: 0.3093 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 1.7483 - sparse_categorical_accuracy: 0.2821 - val_loss: 1.7260 - val_sparse_ca
tegorical_accuracy: 0.3076 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 1.7524 - sparse_categorical_accuracy: 0.2779 - val_loss: 1.7232 - val_sparse_ca
tegorical_accuracy: 0.2951 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 1.7728 - sparse_categorical_accuracy: 0.2803 - val_loss: 1.7572 - val_sparse_ca
tegorical_accuracy: 0.3180 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 1.7653 - sparse_categorical_accuracy: 0.2805 - val_loss: 1.7882 - val_sparse_ca
tegorical_accuracy: 0.3232 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 1.7611 - sparse_categorical_accuracy: 0.2806 - val_loss: 1.7396 - val_sparse_ca
tegorical_accuracy: 0.2806 - 1s/epoch - 3ms/step
Epoch 90/100
469/469 - 1s - loss: 1.7625 - sparse_categorical_accuracy: 0.2829 - val_loss: 1.7272 - val_sparse_ca
tegorical_accuracy: 0.2790 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 1.7311 - sparse_categorical_accuracy: 0.2881 - val_loss: 1.7306 - val_sparse_ca
tegorical_accuracy: 0.2800 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 1.7582 - sparse_categorical_accuracy: 0.2841 - val_loss: 1.7324 - val_sparse_ca
tegorical_accuracy: 0.3052 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 1.7746 - sparse_categorical_accuracy: 0.2812 - val_loss: 1.7492 - val_sparse_ca
tegorical_accuracy: 0.2656 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 1.7541 - sparse_categorical_accuracy: 0.2847 - val_loss: 1.7169 - val_sparse_ca
tegorical_accuracy: 0.3040 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 1.7674 - sparse_categorical_accuracy: 0.2794 - val_loss: 1.7302 - val_sparse_ca
tegorical_accuracy: 0.3114 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 1.7528 - sparse_categorical_accuracy: 0.2855 - val_loss: 1.9863 - val_sparse_ca
tegorical_accuracy: 0.2105 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 1.7842 - sparse_categorical_accuracy: 0.2769 - val_loss: 1.7429 - val_sparse_ca
tegorical_accuracy: 0.2667 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 1.7581 - sparse_categorical_accuracy: 0.2869 - val_loss: 1.7317 - val_sparse_ca
tegorical_accuracy: 0.2647 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 1.7510 - sparse_categorical_accuracy: 0.2846 - val_loss: 1.7315 - val_sparse_ca
tegorical_accuracy: 0.3072 - 1s/epoch - 2ms/step
Epoch 100/100
```

```
469/469 - 1s - loss: 1.7459 - sparse_categorical_accuracy: 0.2835 - val_loss: 1.7336 - val_sparse_ca
tegorical_accuracy: 0.3011 - 1s/epoch - 2ms/step
```

In [106]:

```
model2_f.evaluate(testX_f, testy_f, verbose=2)
```

```
313/313 - 0s - loss: 1.7336 - sparse_categorical_accuracy: 0.3011 - 395ms/epoch - 1ms/step
```

Out[106]:

```
[1.7336406707763672, 0.3010999858379364]
```

Model3: Regularization penalty is L1 with parameter 10^{-3}

In [107]:

```
model3_f = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform',
  activity_regularizer = tf.keras.regularizers.L1(0.001)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
# Fix the learning rate to be 0.001
model3_f.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
history3_f = model3_f.fit(trainX_f, trainy_f,
                     batch_size=128,
                     epochs=100,
                     validation_data=(testX_f, testy_f),
                     verbose=2
                     )
```

```
Epoch 1/100
469/469 - 2s - loss: 1.4526 - sparse_categorical_accuracy: 0.5002 - val_loss: 1.2495 - val_sparse_ca
tegorical_accuracy: 0.6161 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 1.1333 - sparse_categorical_accuracy: 0.6308 - val_loss: 1.0683 - val_sparse_ca
tegorical_accuracy: 0.6436 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 0.9831 - sparse_categorical_accuracy: 0.6835 - val_loss: 0.9379 - val_sparse_ca
tegorical_accuracy: 0.7100 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 0.8937 - sparse_categorical_accuracy: 0.7355 - val_loss: 0.8940 - val_sparse_ca
tegorical_accuracy: 0.7399 - 1s/epoch - 3ms/step
Epoch 5/100
469/469 - 1s - loss: 0.8627 - sparse_categorical_accuracy: 0.7442 - val_loss: 0.8922 - val_sparse_ca
tegorical_accuracy: 0.7419 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 0.8566 - sparse_categorical_accuracy: 0.7486 - val_loss: 0.8978 - val_sparse_ca
tegorical_accuracy: 0.7304 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.8353 - sparse_categorical_accuracy: 0.7556 - val_loss: 0.8594 - val_sparse_ca
tegorical_accuracy: 0.7468 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 0.8358 - sparse_categorical_accuracy: 0.7541 - val_loss: 0.8811 - val_sparse_ca
tegorical_accuracy: 0.7441 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 0.8425 - sparse_categorical_accuracy: 0.7524 - val_loss: 0.8447 - val_sparse_ca
tegorical_accuracy: 0.7565 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 0.8538 - sparse_categorical_accuracy: 0.7513 - val_loss: 0.9363 - val_sparse_ca
tegorical_accuracy: 0.7080 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 0.8541 - sparse_categorical_accuracy: 0.7513 - val_loss: 0.9235 - val_sparse_ca
tegorical_accuracy: 0.7263 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.8542 - sparse_categorical_accuracy: 0.7499 - val_loss: 0.8813 - val_sparse_ca
tegorical_accuracy: 0.7406 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.8632 - sparse_categorical_accuracy: 0.7497 - val_loss: 0.9038 - val_sparse_ca
tegorical_accuracy: 0.7433 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.8417 - sparse_categorical_accuracy: 0.7559 - val_loss: 1.0132 - val_sparse_ca
tegorical_accuracy: 0.6809 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 0.8670 - sparse_categorical_accuracy: 0.7490 - val_loss: 0.8612 - val_sparse_ca
tegorical_accuracy: 0.7530 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.8831 - sparse_categorical_accuracy: 0.7430 - val_loss: 0.9381 - val_sparse_ca
```

```
tegorical_accuracy: 0.7267 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.8631 - sparse_categorical_accuracy: 0.7527 - val_loss: 0.9333 - val_sparse_ca
tegorical_accuracy: 0.7380 - 1s/epoch - 3ms/step
Epoch 18/100
469/469 - 1s - loss: 0.8713 - sparse_categorical_accuracy: 0.7484 - val_loss: 0.8695 - val_sparse_ca
tegorical_accuracy: 0.7624 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.8834 - sparse_categorical_accuracy: 0.7433 - val_loss: 0.8690 - val_sparse_ca
tegorical_accuracy: 0.7518 - 1s/epoch - 3ms/step
Epoch 20/100
469/469 - 1s - loss: 0.8731 - sparse_categorical_accuracy: 0.7523 - val_loss: 0.8926 - val_sparse_ca
tegorical_accuracy: 0.7432 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.8852 - sparse_categorical_accuracy: 0.7472 - val_loss: 0.8835 - val_sparse_ca
tegorical_accuracy: 0.7581 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.8989 - sparse_categorical_accuracy: 0.7435 - val_loss: 0.8987 - val_sparse_ca
tegorical_accuracy: 0.7613 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.8746 - sparse_categorical_accuracy: 0.7557 - val_loss: 0.9290 - val_sparse_ca
tegorical_accuracy: 0.7378 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.8869 - sparse_categorical_accuracy: 0.7488 - val_loss: 1.0127 - val_sparse_ca
tegorical_accuracy: 0.7205 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.9082 - sparse_categorical_accuracy: 0.7408 - val_loss: 0.9394 - val_sparse_ca
tegorical_accuracy: 0.7323 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.9240 - sparse_categorical_accuracy: 0.7366 - val_loss: 0.9898 - val_sparse_ca
tegorical_accuracy: 0.7349 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.9212 - sparse_categorical_accuracy: 0.7451 - val_loss: 0.9251 - val_sparse_ca
tegorical_accuracy: 0.7345 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.9020 - sparse_categorical_accuracy: 0.7471 - val_loss: 0.8922 - val_sparse_ca
tegorical_accuracy: 0.7553 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.8944 - sparse_categorical_accuracy: 0.7410 - val_loss: 0.8956 - val_sparse_ca
tegorical_accuracy: 0.7401 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.9158 - sparse_categorical_accuracy: 0.7457 - val_loss: 1.2868 - val_sparse_ca
tegorical_accuracy: 0.6734 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.9454 - sparse_categorical_accuracy: 0.7395 - val_loss: 0.9442 - val_sparse_ca
tegorical_accuracy: 0.7279 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.9400 - sparse_categorical_accuracy: 0.7425 - val_loss: 1.2926 - val_sparse_ca
tegorical_accuracy: 0.6724 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.9130 - sparse_categorical_accuracy: 0.7493 - val_loss: 0.9316 - val_sparse_ca
tegorical_accuracy: 0.7431 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.9028 - sparse_categorical_accuracy: 0.7496 - val_loss: 1.0032 - val_sparse_ca
tegorical_accuracy: 0.7144 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.9818 - sparse_categorical_accuracy: 0.7245 - val_loss: 0.9520 - val_sparse_ca
tegorical_accuracy: 0.7415 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.9807 - sparse_categorical_accuracy: 0.7346 - val_loss: 1.0927 - val_sparse_ca
tegorical_accuracy: 0.7182 - 1s/epoch - 3ms/step
Epoch 37/100
469/469 - 1s - loss: 0.9263 - sparse_categorical_accuracy: 0.7484 - val_loss: 0.9142 - val_sparse_ca
tegorical_accuracy: 0.7518 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.9537 - sparse_categorical_accuracy: 0.7422 - val_loss: 0.9502 - val_sparse_ca
tegorical_accuracy: 0.7468 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.9466 - sparse_categorical_accuracy: 0.7401 - val_loss: 1.0285 - val_sparse_ca
tegorical_accuracy: 0.7287 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 0.9709 - sparse_categorical_accuracy: 0.7326 - val_loss: 0.9428 - val_sparse_ca
tegorical_accuracy: 0.7492 - 1s/epoch - 3ms/step
Epoch 41/100
469/469 - 1s - loss: 0.9389 - sparse_categorical_accuracy: 0.7469 - val_loss: 1.0392 - val_sparse_ca
tegorical_accuracy: 0.7285 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.9418 - sparse_categorical_accuracy: 0.7450 - val_loss: 1.0152 - val_sparse_ca
tegorical_accuracy: 0.7286 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.9690 - sparse_categorical_accuracy: 0.7268 - val_loss: 0.9498 - val_sparse_ca
tegorical_accuracy: 0.7350 - 1s/epoch - 2ms/step
Epoch 44/100
```

```
469/469 - 1s - loss: 0.9640 - sparse_categorical_accuracy: 0.7408 - val_loss: 1.1167 - val_sparse_ca
tegorical_accuracy: 0.6785 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.9915 - sparse_categorical_accuracy: 0.7358 - val_loss: 0.9822 - val_sparse_ca
tegorical_accuracy: 0.7363 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.9769 - sparse_categorical_accuracy: 0.7424 - val_loss: 1.0120 - val_sparse_ca
tegorical_accuracy: 0.7205 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.9595 - sparse_categorical_accuracy: 0.7408 - val_loss: 0.9507 - val_sparse_ca
tegorical_accuracy: 0.7515 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.9692 - sparse_categorical_accuracy: 0.7393 - val_loss: 1.0260 - val_sparse_ca
tegorical_accuracy: 0.7218 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.9862 - sparse_categorical_accuracy: 0.7401 - val_loss: 1.0162 - val_sparse_ca
tegorical_accuracy: 0.7285 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.9604 - sparse_categorical_accuracy: 0.7388 - val_loss: 0.9465 - val_sparse_ca
tegorical_accuracy: 0.7506 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.9763 - sparse_categorical_accuracy: 0.7405 - val_loss: 0.9648 - val_sparse_ca
tegorical_accuracy: 0.7360 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.9795 - sparse_categorical_accuracy: 0.7395 - val_loss: 1.0772 - val_sparse_ca
tegorical_accuracy: 0.7285 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.9721 - sparse_categorical_accuracy: 0.7409 - val_loss: 0.9726 - val_sparse_ca
tegorical_accuracy: 0.7492 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 1.0106 - sparse_categorical_accuracy: 0.7299 - val_loss: 1.0098 - val_sparse_ca
tegorical_accuracy: 0.7177 - 1s/epoch - 3ms/step
Epoch 55/100
469/469 - 1s - loss: 0.9480 - sparse_categorical_accuracy: 0.7462 - val_loss: 0.9704 - val_sparse_ca
tegorical_accuracy: 0.7484 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.9730 - sparse_categorical_accuracy: 0.7434 - val_loss: 1.1032 - val_sparse_ca
tegorical_accuracy: 0.7002 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 0.9785 - sparse_categorical_accuracy: 0.7456 - val_loss: 0.9958 - val_sparse_ca
tegorical_accuracy: 0.7250 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.9489 - sparse_categorical_accuracy: 0.7361 - val_loss: 0.9420 - val_sparse_ca
tegorical_accuracy: 0.7362 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 1.0301 - sparse_categorical_accuracy: 0.7223 - val_loss: 1.0235 - val_sparse_ca
tegorical_accuracy: 0.7470 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 1.1438 - sparse_categorical_accuracy: 0.6319 - val_loss: 1.1334 - val_sparse_ca
tegorical_accuracy: 0.6263 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 1.0775 - sparse_categorical_accuracy: 0.7014 - val_loss: 1.1998 - val_sparse_ca
tegorical_accuracy: 0.6724 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.9937 - sparse_categorical_accuracy: 0.7355 - val_loss: 1.0905 - val_sparse_ca
tegorical_accuracy: 0.6925 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 1.0529 - sparse_categorical_accuracy: 0.7302 - val_loss: 1.5242 - val_sparse_ca
tegorical_accuracy: 0.6315 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 1.0506 - sparse_categorical_accuracy: 0.7196 - val_loss: 1.1533 - val_sparse_ca
tegorical_accuracy: 0.6658 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.9973 - sparse_categorical_accuracy: 0.7379 - val_loss: 0.9610 - val_sparse_ca
tegorical_accuracy: 0.7459 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.9735 - sparse_categorical_accuracy: 0.7380 - val_loss: 0.9947 - val_sparse_ca
tegorical_accuracy: 0.7182 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.9807 - sparse_categorical_accuracy: 0.7397 - val_loss: 1.0053 - val_sparse_ca
tegorical_accuracy: 0.7318 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 0.9970 - sparse_categorical_accuracy: 0.7321 - val_loss: 1.3808 - val_sparse_ca
tegorical_accuracy: 0.6423 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 1.0973 - sparse_categorical_accuracy: 0.7172 - val_loss: 1.0480 - val_sparse_ca
tegorical_accuracy: 0.7359 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.9954 - sparse_categorical_accuracy: 0.7413 - val_loss: 0.9905 - val_sparse_ca
tegorical_accuracy: 0.7526 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 1.0219 - sparse_categorical_accuracy: 0.7289 - val_loss: 1.0910 - val_sparse_ca
tegorical_accuracy: 0.7230 - 1s/epoch - 2ms/step
```

```
Epoch 72/100
469/469 - 1s - loss: 1.0042 - sparse_categorical_accuracy: 0.7191 - val_loss: 1.0002 - val_sparse_ca
tegorical_accuracy: 0.7284 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 1.1209 - sparse_categorical_accuracy: 0.7128 - val_loss: 1.0790 - val_sparse_ca
tegorical_accuracy: 0.7259 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 1.0242 - sparse_categorical_accuracy: 0.7254 - val_loss: 1.0620 - val_sparse_ca
tegorical_accuracy: 0.7255 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 1.0476 - sparse_categorical_accuracy: 0.7245 - val_loss: 1.0497 - val_sparse_ca
tegorical_accuracy: 0.7360 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 1.0430 - sparse_categorical_accuracy: 0.7336 - val_loss: 1.4391 - val_sparse_ca
tegorical_accuracy: 0.6245 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 1.0312 - sparse_categorical_accuracy: 0.7374 - val_loss: 0.9589 - val_sparse_ca
tegorical_accuracy: 0.7425 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.9938 - sparse_categorical_accuracy: 0.7445 - val_loss: 0.9810 - val_sparse_ca
tegorical_accuracy: 0.7538 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 1.0902 - sparse_categorical_accuracy: 0.6926 - val_loss: 1.1958 - val_sparse_ca
tegorical_accuracy: 0.6593 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 1.0333 - sparse_categorical_accuracy: 0.7277 - val_loss: 1.0398 - val_sparse_ca
tegorical_accuracy: 0.7035 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 1.0239 - sparse_categorical_accuracy: 0.7361 - val_loss: 1.4232 - val_sparse_ca
tegorical_accuracy: 0.6394 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 1.0764 - sparse_categorical_accuracy: 0.7337 - val_loss: 1.0350 - val_sparse_ca
tegorical_accuracy: 0.7218 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 1.0331 - sparse_categorical_accuracy: 0.7315 - val_loss: 1.0321 - val_sparse_ca
tegorical_accuracy: 0.7301 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 1.0656 - sparse_categorical_accuracy: 0.7107 - val_loss: 1.2963 - val_sparse_ca
tegorical_accuracy: 0.5506 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 1.1276 - sparse_categorical_accuracy: 0.6230 - val_loss: 1.1198 - val_sparse_ca
tegorical_accuracy: 0.6264 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 1.0393 - sparse_categorical_accuracy: 0.6739 - val_loss: 1.0888 - val_sparse_ca
tegorical_accuracy: 0.6490 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 1.0470 - sparse_categorical_accuracy: 0.6834 - val_loss: 1.0206 - val_sparse_ca
tegorical_accuracy: 0.6792 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 1.0274 - sparse_categorical_accuracy: 0.6967 - val_loss: 1.0567 - val_sparse_ca
tegorical_accuracy: 0.7169 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 1.0300 - sparse_categorical_accuracy: 0.6915 - val_loss: 1.1531 - val_sparse_ca
tegorical_accuracy: 0.6673 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 1.0835 - sparse_categorical_accuracy: 0.6408 - val_loss: 1.0825 - val_sparse_ca
tegorical_accuracy: 0.6743 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 1.0475 - sparse_categorical_accuracy: 0.6747 - val_loss: 1.0177 - val_sparse_ca
tegorical_accuracy: 0.6795 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 1.1211 - sparse_categorical_accuracy: 0.6389 - val_loss: 1.3113 - val_sparse_ca
tegorical_accuracy: 0.5354 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 1.2200 - sparse_categorical_accuracy: 0.5669 - val_loss: 1.2446 - val_sparse_ca
tegorical_accuracy: 0.5860 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 1.1882 - sparse_categorical_accuracy: 0.5903 - val_loss: 1.2358 - val_sparse_ca
tegorical_accuracy: 0.5757 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 1.2167 - sparse_categorical_accuracy: 0.5756 - val_loss: 1.2448 - val_sparse_ca
tegorical_accuracy: 0.5760 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 1.2061 - sparse_categorical_accuracy: 0.5881 - val_loss: 1.2409 - val_sparse_ca
tegorical_accuracy: 0.5731 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 1.2172 - sparse_categorical_accuracy: 0.5564 - val_loss: 1.1920 - val_sparse_ca
tegorical_accuracy: 0.6002 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 1.1944 - sparse_categorical_accuracy: 0.5978 - val_loss: 1.2069 - val_sparse_ca
tegorical_accuracy: 0.6133 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 1.2230 - sparse_categorical_accuracy: 0.5968 - val_loss: 1.2132 - val_sparse_ca
```

```
tegorical_accuracy: 0.5769 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 1.3083 - sparse_categorical_accuracy: 0.5722 - val_loss: 1.2296 - val_sparse_ca
tegorical_accuracy: 0.5905 - 1s/epoch - 2ms/step
```

In [108]:

```
model3_f.evaluate(testX_f, testy_f, verbose=2)
```

```
313/313 - 0s - loss: 1.2296 - sparse_categorical_accuracy: 0.5905 - 391ms/epoch - 1ms/step
```

Out[108]:

```
[1.229557752609253, 0.590499997138977]
```

Model4: Regularization penalty is L2 with parameter 10^{-2}

In [109]:

```
model4_f = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform',
  activity_regularizer = tf.keras.regularizers.L2(0.01)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
# Fix the learning rate to be 0.001
model4_f.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
history4_f = model4_f.fit(trainX_f, trainy_f,
                     batch_size=128,
                     epochs=100,
                     validation_data=(testX_f, testy_f),
                     verbose=2
                     )
```

```
Epoch 1/100
469/469 - 2s - loss: 4.3971 - sparse_categorical_accuracy: 0.0969 - val_loss: 2.3030 - val_sparse_ca
tegorical_accuracy: 0.1002 - 2s/epoch - 5ms/step
Epoch 2/100
469/469 - 1s - loss: 2.3029 - sparse_categorical_accuracy: 0.0974 - val_loss: 2.3029 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 2.3028 - sparse_categorical_accuracy: 0.0980 - val_loss: 2.3029 - val_sparse_ca
tegorical_accuracy: 0.1002 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 2.3028 - sparse_categorical_accuracy: 0.0990 - val_loss: 2.3029 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 2.3028 - sparse_categorical_accuracy: 0.0981 - val_loss: 2.3029 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0976 - val_loss: 2.3029 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3029 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0972 - val_loss: 2.3029 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3029 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0978 - val_loss: 2.3029 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0995 - val_loss: 2.3029 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0981 - val_loss: 2.3029 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.1000 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 2.3028 - sparse_categorical_accuracy: 0.0966 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1002 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0991 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
```

```
Epoch 16/100
469/469 - 1s - loss: 2.3029 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 2.3026 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0991 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0983 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0988 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0974 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0967 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0975 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0976 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0967 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0979 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0981 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0983 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0972 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0989 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0988 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0979 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0993 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0990 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0979 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0971 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0976 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0974 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0972 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0984 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0991 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0988 - val_loss: 2.3028 - val_sparse_ca
```

```
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0989 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0988 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0989 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.1001 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0972 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0988 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0973 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0983 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0964 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0995 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0990 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0975 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0966 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0982 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0992 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0977 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0974 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0990 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0961 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0983 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0989 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0979 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.1001 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0986 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 71/100
```

```
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0983 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0983 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0968 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0987 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0989 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0992 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0970 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0990 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0982 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0988 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0984 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0962 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0983 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0966 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0980 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0976 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0973 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 90/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0980 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0977 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0970 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0985 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0962 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 95/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0990 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0974 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1000 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0973 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 3ms/step
Epoch 98/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0991 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
```

```
Epoch 99/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0986 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 2.3027 - sparse_categorical_accuracy: 0.0994 - val_loss: 2.3028 - val_sparse_ca
tegorical_accuracy: 0.1001 - 1s/epoch - 2ms/step
```

In [110]:

```
model4_f.evaluate(testX_f, testy_f, verbose=2)
```

```
313/313 - 0s - loss: 2.3028 - sparse_categorical_accuracy: 0.1001 - 399ms/epoch - 1ms/step
```

Out[110]:

```
[2.3028159141540527, 0.10010000318288803]
```

Model5: Regularization penalty is L2 with parameter 10^{-4}

In [111]:

```
model5_f = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform',
  activity_regularizer = tf.keras.regularizers.L2(0.0001)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
# Fix the learning rate to be 0.001
model5_f.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
history5_f = model5_f.fit(trainX_f, trainy_f,
                     batch_size=128,
                     epochs=100,
                     validation_data=(testX_f, testy_f),
                     verbose=2
                     )
```

```
Epoch 1/100
469/469 - 2s - loss: 2.1561 - sparse_categorical_accuracy: 0.1645 - val_loss: 1.9091 - val_sparse_ca
tegorical_accuracy: 0.2218 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 1.8159 - sparse_categorical_accuracy: 0.2522 - val_loss: 1.8304 - val_sparse_ca
tegorical_accuracy: 0.2576 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 1.7879 - sparse_categorical_accuracy: 0.2719 - val_loss: 1.7549 - val_sparse_ca
tegorical_accuracy: 0.2780 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 1.7232 - sparse_categorical_accuracy: 0.2996 - val_loss: 1.7395 - val_sparse_ca
tegorical_accuracy: 0.2709 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 1.7257 - sparse_categorical_accuracy: 0.3060 - val_loss: 1.6495 - val_sparse_ca
tegorical_accuracy: 0.3458 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 1.6989 - sparse_categorical_accuracy: 0.3144 - val_loss: 1.6691 - val_sparse_ca
tegorical_accuracy: 0.3331 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 1.7378 - sparse_categorical_accuracy: 0.3139 - val_loss: 1.6540 - val_sparse_ca
tegorical_accuracy: 0.3212 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 1.6877 - sparse_categorical_accuracy: 0.3203 - val_loss: 1.7275 - val_sparse_ca
tegorical_accuracy: 0.2900 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 1.6823 - sparse_categorical_accuracy: 0.3258 - val_loss: 1.6540 - val_sparse_ca
tegorical_accuracy: 0.3082 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 1.7364 - sparse_categorical_accuracy: 0.3010 - val_loss: 1.8741 - val_sparse_ca
tegorical_accuracy: 0.2659 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 1.6907 - sparse_categorical_accuracy: 0.3288 - val_loss: 1.6507 - val_sparse_ca
tegorical_accuracy: 0.3261 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 1.7167 - sparse_categorical_accuracy: 0.3131 - val_loss: 1.6424 - val_sparse_ca
tegorical_accuracy: 0.3423 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 1.7119 - sparse_categorical_accuracy: 0.3213 - val_loss: 1.7571 - val_sparse_ca
tegorical_accuracy: 0.2874 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 1.7354 - sparse_categorical_accuracy: 0.3140 - val_loss: 1.6449 - val_sparse_ca
tegorical_accuracy: 0.3439 - 1s/epoch - 2ms/step
Epoch 15/100
```

```
469/469 - 1s - loss: 1.6982 - sparse_categorical_accuracy: 0.3216 - val_loss: 1.6751 - val_sparse_ca
tegorical_accuracy: 0.3187 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 1.7373 - sparse_categorical_accuracy: 0.3168 - val_loss: 1.6850 - val_sparse_ca
tegorical_accuracy: 0.3007 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 1.7567 - sparse_categorical_accuracy: 0.3122 - val_loss: 1.6879 - val_sparse_ca
tegorical_accuracy: 0.3268 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 1.7756 - sparse_categorical_accuracy: 0.3115 - val_loss: 1.7339 - val_sparse_ca
tegorical_accuracy: 0.3536 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 1.7960 - sparse_categorical_accuracy: 0.3061 - val_loss: 1.7862 - val_sparse_ca
tegorical_accuracy: 0.3030 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 1.7430 - sparse_categorical_accuracy: 0.3099 - val_loss: 1.8425 - val_sparse_ca
tegorical_accuracy: 0.2740 - 1s/epoch - 3ms/step
Epoch 21/100
469/469 - 1s - loss: 1.8447 - sparse_categorical_accuracy: 0.2946 - val_loss: 1.8784 - val_sparse_ca
tegorical_accuracy: 0.3510 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 1.8329 - sparse_categorical_accuracy: 0.2950 - val_loss: 1.8328 - val_sparse_ca
tegorical_accuracy: 0.3222 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 1.7851 - sparse_categorical_accuracy: 0.3008 - val_loss: 1.8746 - val_sparse_ca
tegorical_accuracy: 0.2738 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 1.8235 - sparse_categorical_accuracy: 0.2875 - val_loss: 1.7645 - val_sparse_ca
tegorical_accuracy: 0.3083 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 1.8198 - sparse_categorical_accuracy: 0.2878 - val_loss: 1.7652 - val_sparse_ca
tegorical_accuracy: 0.2621 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 1.8299 - sparse_categorical_accuracy: 0.2963 - val_loss: 1.9277 - val_sparse_ca
tegorical_accuracy: 0.2684 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 1.8204 - sparse_categorical_accuracy: 0.2907 - val_loss: 2.1030 - val_sparse_ca
tegorical_accuracy: 0.3346 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 1.8428 - sparse_categorical_accuracy: 0.2857 - val_loss: 1.7631 - val_sparse_ca
tegorical_accuracy: 0.2640 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 1.9763 - sparse_categorical_accuracy: 0.2582 - val_loss: 1.8495 - val_sparse_ca
tegorical_accuracy: 0.2641 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 1.9863 - sparse_categorical_accuracy: 0.2601 - val_loss: 1.8445 - val_sparse_ca
tegorical_accuracy: 0.2490 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 1.9448 - sparse_categorical_accuracy: 0.2539 - val_loss: 1.8970 - val_sparse_ca
tegorical_accuracy: 0.2572 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 1.9060 - sparse_categorical_accuracy: 0.2645 - val_loss: 1.8848 - val_sparse_ca
tegorical_accuracy: 0.3005 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 1.8620 - sparse_categorical_accuracy: 0.2765 - val_loss: 1.8446 - val_sparse_ca
tegorical_accuracy: 0.2566 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 1.9507 - sparse_categorical_accuracy: 0.2587 - val_loss: 1.8481 - val_sparse_ca
tegorical_accuracy: 0.2860 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 1.9673 - sparse_categorical_accuracy: 0.2506 - val_loss: 2.1567 - val_sparse_ca
tegorical_accuracy: 0.1776 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 1.9739 - sparse_categorical_accuracy: 0.2539 - val_loss: 1.8668 - val_sparse_ca
tegorical_accuracy: 0.2579 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 1.9576 - sparse_categorical_accuracy: 0.2550 - val_loss: 1.9598 - val_sparse_ca
tegorical_accuracy: 0.2434 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 2.0848 - sparse_categorical_accuracy: 0.2346 - val_loss: 1.9605 - val_sparse_ca
tegorical_accuracy: 0.2518 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 2.0129 - sparse_categorical_accuracy: 0.2492 - val_loss: 1.9316 - val_sparse_ca
tegorical_accuracy: 0.2565 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 2.0596 - sparse_categorical_accuracy: 0.2319 - val_loss: 2.1240 - val_sparse_ca
tegorical_accuracy: 0.1974 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 2.0222 - sparse_categorical_accuracy: 0.2390 - val_loss: 2.5750 - val_sparse_ca
tegorical_accuracy: 0.2518 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 2.0781 - sparse_categorical_accuracy: 0.2317 - val_loss: 2.0979 - val_sparse_ca
tegorical_accuracy: 0.1865 - 1s/epoch - 2ms/step
```

```
Epoch 43/100
469/469 - 1s - loss: 2.1911 - sparse_categorical_accuracy: 0.1762 - val_loss: 2.2162 - val_sparse_ca
tegorical_accuracy: 0.1414 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 2.2203 - sparse_categorical_accuracy: 0.1509 - val_loss: 2.2054 - val_sparse_ca
tegorical_accuracy: 0.1485 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 2.2209 - sparse_categorical_accuracy: 0.1528 - val_loss: 2.2281 - val_sparse_ca
tegorical_accuracy: 0.1415 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 2.2186 - sparse_categorical_accuracy: 0.1582 - val_loss: 2.1912 - val_sparse_ca
tegorical_accuracy: 0.1655 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 2.2488 - sparse_categorical_accuracy: 0.1629 - val_loss: 2.1375 - val_sparse_ca
tegorical_accuracy: 0.2252 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 2.1630 - sparse_categorical_accuracy: 0.2016 - val_loss: 2.1070 - val_sparse_ca
tegorical_accuracy: 0.2213 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 2.1870 - sparse_categorical_accuracy: 0.2001 - val_loss: 2.1692 - val_sparse_ca
tegorical_accuracy: 0.1786 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 2.1418 - sparse_categorical_accuracy: 0.2139 - val_loss: 2.1030 - val_sparse_ca
tegorical_accuracy: 0.1988 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 2.1134 - sparse_categorical_accuracy: 0.1996 - val_loss: 2.1029 - val_sparse_ca
tegorical_accuracy: 0.1936 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 2.1467 - sparse_categorical_accuracy: 0.2044 - val_loss: 2.1190 - val_sparse_ca
tegorical_accuracy: 0.1925 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 2.1308 - sparse_categorical_accuracy: 0.2077 - val_loss: 2.0666 - val_sparse_ca
tegorical_accuracy: 0.2360 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 2.1760 - sparse_categorical_accuracy: 0.2103 - val_loss: 2.0649 - val_sparse_ca
tegorical_accuracy: 0.2427 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 2.0936 - sparse_categorical_accuracy: 0.2156 - val_loss: 2.1255 - val_sparse_ca
tegorical_accuracy: 0.1900 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 2.1386 - sparse_categorical_accuracy: 0.2214 - val_loss: 2.1239 - val_sparse_ca
tegorical_accuracy: 0.2137 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 2.1831 - sparse_categorical_accuracy: 0.2204 - val_loss: 2.0423 - val_sparse_ca
tegorical_accuracy: 0.2367 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 2.1529 - sparse_categorical_accuracy: 0.2143 - val_loss: 2.1986 - val_sparse_ca
tegorical_accuracy: 0.1603 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 2.1346 - sparse_categorical_accuracy: 0.1973 - val_loss: 2.0997 - val_sparse_ca
tegorical_accuracy: 0.2052 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 2.1305 - sparse_categorical_accuracy: 0.2114 - val_loss: 2.1033 - val_sparse_ca
tegorical_accuracy: 0.2118 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 2.1719 - sparse_categorical_accuracy: 0.2096 - val_loss: 2.1113 - val_sparse_ca
tegorical_accuracy: 0.2126 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 2.1584 - sparse_categorical_accuracy: 0.2063 - val_loss: 2.0850 - val_sparse_ca
tegorical_accuracy: 0.2171 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 2.1175 - sparse_categorical_accuracy: 0.2115 - val_loss: 2.0669 - val_sparse_ca
tegorical_accuracy: 0.2174 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 2.1108 - sparse_categorical_accuracy: 0.2145 - val_loss: 2.0977 - val_sparse_ca
tegorical_accuracy: 0.2019 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 2.1398 - sparse_categorical_accuracy: 0.1992 - val_loss: 2.1610 - val_sparse_ca
tegorical_accuracy: 0.1690 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 2.1994 - sparse_categorical_accuracy: 0.2080 - val_loss: 2.1568 - val_sparse_ca
tegorical_accuracy: 0.2171 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 2.1668 - sparse_categorical_accuracy: 0.1992 - val_loss: 2.1224 - val_sparse_ca
tegorical_accuracy: 0.2093 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 2.1586 - sparse_categorical_accuracy: 0.1896 - val_loss: 2.1215 - val_sparse_ca
tegorical_accuracy: 0.1965 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 2.1741 - sparse_categorical_accuracy: 0.1967 - val_loss: 2.1922 - val_sparse_ca
tegorical_accuracy: 0.1802 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 2.1731 - sparse_categorical_accuracy: 0.1936 - val_loss: 2.1910 - val_sparse_ca
```

```
tegorical_accuracy: 0.2135 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 2.2206 - sparse_categorical_accuracy: 0.1922 - val_loss: 2.1927 - val_sparse_ca
tegorical_accuracy: 0.1681 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 2.2165 - sparse_categorical_accuracy: 0.1864 - val_loss: 2.1877 - val_sparse_ca
tegorical_accuracy: 0.2177 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 2.1961 - sparse_categorical_accuracy: 0.1977 - val_loss: 2.2139 - val_sparse_ca
tegorical_accuracy: 0.1969 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 2.1633 - sparse_categorical_accuracy: 0.1910 - val_loss: 2.1671 - val_sparse_ca
tegorical_accuracy: 0.1816 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 2.2121 - sparse_categorical_accuracy: 0.2032 - val_loss: 2.1897 - val_sparse_ca
tegorical_accuracy: 0.1632 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 2.1514 - sparse_categorical_accuracy: 0.1971 - val_loss: 2.1698 - val_sparse_ca
tegorical_accuracy: 0.1679 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 2.1356 - sparse_categorical_accuracy: 0.1959 - val_loss: 2.2192 - val_sparse_ca
tegorical_accuracy: 0.1549 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 2.2772 - sparse_categorical_accuracy: 0.1535 - val_loss: 2.2479 - val_sparse_ca
tegorical_accuracy: 0.1321 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 2.2508 - sparse_categorical_accuracy: 0.1529 - val_loss: 2.5524 - val_sparse_ca
tegorical_accuracy: 0.2105 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 2.2218 - sparse_categorical_accuracy: 0.1887 - val_loss: 2.1746 - val_sparse_ca
tegorical_accuracy: 0.1665 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 2.2122 - sparse_categorical_accuracy: 0.1793 - val_loss: 2.4575 - val_sparse_ca
tegorical_accuracy: 0.2062 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 2.1881 - sparse_categorical_accuracy: 0.1906 - val_loss: 2.1955 - val_sparse_ca
tegorical_accuracy: 0.2166 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 2.1922 - sparse_categorical_accuracy: 0.1728 - val_loss: 2.1820 - val_sparse_ca
tegorical_accuracy: 0.1568 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 2.2389 - sparse_categorical_accuracy: 0.1723 - val_loss: 2.2216 - val_sparse_ca
tegorical_accuracy: 0.1992 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 2.2081 - sparse_categorical_accuracy: 0.1832 - val_loss: 2.1244 - val_sparse_ca
tegorical_accuracy: 0.1903 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 2.1932 - sparse_categorical_accuracy: 0.1899 - val_loss: 2.5147 - val_sparse_ca
tegorical_accuracy: 0.2133 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 2.1917 - sparse_categorical_accuracy: 0.1905 - val_loss: 2.1390 - val_sparse_ca
tegorical_accuracy: 0.1788 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 2.1964 - sparse_categorical_accuracy: 0.1840 - val_loss: 2.1836 - val_sparse_ca
tegorical_accuracy: 0.1842 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 2.2389 - sparse_categorical_accuracy: 0.1784 - val_loss: 2.1511 - val_sparse_ca
tegorical_accuracy: 0.1746 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 2.2384 - sparse_categorical_accuracy: 0.1792 - val_loss: 2.2464 - val_sparse_ca
tegorical_accuracy: 0.1317 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 2.2319 - sparse_categorical_accuracy: 0.1713 - val_loss: 2.1824 - val_sparse_ca
tegorical_accuracy: 0.1749 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 2.1929 - sparse_categorical_accuracy: 0.1706 - val_loss: 2.1661 - val_sparse_ca
tegorical_accuracy: 0.1664 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 2.2160 - sparse_categorical_accuracy: 0.1599 - val_loss: 2.1760 - val_sparse_ca
tegorical_accuracy: 0.1678 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 2.2165 - sparse_categorical_accuracy: 0.1603 - val_loss: 2.3475 - val_sparse_ca
tegorical_accuracy: 0.1983 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 2.2274 - sparse_categorical_accuracy: 0.1661 - val_loss: 2.1990 - val_sparse_ca
tegorical_accuracy: 0.1580 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 2.2658 - sparse_categorical_accuracy: 0.1286 - val_loss: 2.2951 - val_sparse_ca
tegorical_accuracy: 0.1326 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 2.2851 - sparse_categorical_accuracy: 0.1119 - val_loss: 2.2922 - val_sparse_ca
tegorical_accuracy: 0.1066 - 1s/epoch - 2ms/step
Epoch 98/100
```

```
469/469 - 1s - loss: 2.2835 - sparse_categorical_accuracy: 0.1170 - val_loss: 2.2867 - val_sparse_ca
tegorical_accuracy: 0.1103 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 2.2902 - sparse_categorical_accuracy: 0.1178 - val_loss: 2.2888 - val_sparse_ca
tegorical_accuracy: 0.1139 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 2.2836 - sparse_categorical_accuracy: 0.1142 - val_loss: 2.2809 - val_sparse_ca
tegorical_accuracy: 0.1180 - 1s/epoch - 2ms/step
```

In [112]:

```
model5_f.evaluate(testX_f, testy_f, verbose=2)
```

```
313/313 - 0s - loss: 2.2809 - sparse_categorical_accuracy: 0.1180 - 398ms/epoch - 1ms/step
```

Out[112]:

```
[2.2808923721313477, 0.11800000071525574]
```

Model6: Regularization penalty is L2 with parameter 10^{-5}

In [113]:

```
model6_f = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform',
  activity_regularizer = tf.keras.regularizers.L2(0.00001)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
# Fix the learning rate to be 0.001
model6_f.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
history6_f = model6_f.fit(trainX_f, trainy_f,
                      batch_size=128,
                      epochs=100,
                      validation_data=(testX_f, testy_f),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 2s - loss: 1.3710 - sparse_categorical_accuracy: 0.5325 - val_loss: 1.0294 - val_sparse_ca
tegorical_accuracy: 0.6598 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 0.9311 - sparse_categorical_accuracy: 0.6939 - val_loss: 0.8892 - val_sparse_ca
tegorical_accuracy: 0.7159 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 0.8112 - sparse_categorical_accuracy: 0.7189 - val_loss: 0.8756 - val_sparse_ca
tegorical_accuracy: 0.7154 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 0.7896 - sparse_categorical_accuracy: 0.7200 - val_loss: 0.8679 - val_sparse_ca
tegorical_accuracy: 0.7123 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 0.7755 - sparse_categorical_accuracy: 0.7232 - val_loss: 0.8608 - val_sparse_ca
tegorical_accuracy: 0.7130 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 0.7696 - sparse_categorical_accuracy: 0.7242 - val_loss: 0.8317 - val_sparse_ca
tegorical_accuracy: 0.7118 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.7791 - sparse_categorical_accuracy: 0.7265 - val_loss: 0.7673 - val_sparse_ca
tegorical_accuracy: 0.7260 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 0.7665 - sparse_categorical_accuracy: 0.7252 - val_loss: 0.7821 - val_sparse_ca
tegorical_accuracy: 0.7141 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 0.7752 - sparse_categorical_accuracy: 0.7222 - val_loss: 0.8061 - val_sparse_ca
tegorical_accuracy: 0.7048 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 0.7741 - sparse_categorical_accuracy: 0.7250 - val_loss: 0.8258 - val_sparse_ca
tegorical_accuracy: 0.7110 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 0.7768 - sparse_categorical_accuracy: 0.7233 - val_loss: 0.8286 - val_sparse_ca
tegorical_accuracy: 0.7102 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.7891 - sparse_categorical_accuracy: 0.7212 - val_loss: 0.8288 - val_sparse_ca
tegorical_accuracy: 0.7069 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.7790 - sparse_categorical_accuracy: 0.7218 - val_loss: 0.7924 - val_sparse_ca
tegorical_accuracy: 0.7168 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.8076 - sparse_categorical_accuracy: 0.7197 - val_loss: 0.8958 - val_sparse_ca
```

```
tegorical_accuracy: 0.7184 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 0.7897 - sparse_categorical_accuracy: 0.7224 - val_loss: 0.9659 - val_sparse_ca
tegorical_accuracy: 0.7215 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.8015 - sparse_categorical_accuracy: 0.7232 - val_loss: 0.9060 - val_sparse_ca
tegorical_accuracy: 0.6929 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.7906 - sparse_categorical_accuracy: 0.7234 - val_loss: 0.8166 - val_sparse_ca
tegorical_accuracy: 0.7239 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 0.8323 - sparse_categorical_accuracy: 0.7199 - val_loss: 0.8686 - val_sparse_ca
tegorical_accuracy: 0.7018 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.7985 - sparse_categorical_accuracy: 0.7241 - val_loss: 0.8116 - val_sparse_ca
tegorical_accuracy: 0.7028 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.8273 - sparse_categorical_accuracy: 0.7198 - val_loss: 0.8147 - val_sparse_ca
tegorical_accuracy: 0.7194 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.8126 - sparse_categorical_accuracy: 0.7198 - val_loss: 0.8053 - val_sparse_ca
tegorical_accuracy: 0.7224 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.8302 - sparse_categorical_accuracy: 0.7159 - val_loss: 0.8417 - val_sparse_ca
tegorical_accuracy: 0.7123 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.8428 - sparse_categorical_accuracy: 0.7158 - val_loss: 0.9623 - val_sparse_ca
tegorical_accuracy: 0.7207 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.8403 - sparse_categorical_accuracy: 0.7190 - val_loss: 0.9864 - val_sparse_ca
tegorical_accuracy: 0.7055 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.8521 - sparse_categorical_accuracy: 0.7141 - val_loss: 0.8857 - val_sparse_ca
tegorical_accuracy: 0.7109 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.8950 - sparse_categorical_accuracy: 0.7141 - val_loss: 0.8567 - val_sparse_ca
tegorical_accuracy: 0.7166 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.8681 - sparse_categorical_accuracy: 0.7164 - val_loss: 0.9699 - val_sparse_ca
tegorical_accuracy: 0.6891 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.8931 - sparse_categorical_accuracy: 0.7140 - val_loss: 1.0250 - val_sparse_ca
tegorical_accuracy: 0.6653 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.9067 - sparse_categorical_accuracy: 0.7053 - val_loss: 0.9987 - val_sparse_ca
tegorical_accuracy: 0.6652 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.9206 - sparse_categorical_accuracy: 0.7043 - val_loss: 0.9914 - val_sparse_ca
tegorical_accuracy: 0.6973 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.9050 - sparse_categorical_accuracy: 0.7051 - val_loss: 0.8455 - val_sparse_ca
tegorical_accuracy: 0.7063 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.8991 - sparse_categorical_accuracy: 0.7070 - val_loss: 0.9170 - val_sparse_ca
tegorical_accuracy: 0.7126 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.9367 - sparse_categorical_accuracy: 0.7031 - val_loss: 0.9097 - val_sparse_ca
tegorical_accuracy: 0.7145 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.9274 - sparse_categorical_accuracy: 0.7099 - val_loss: 0.9642 - val_sparse_ca
tegorical_accuracy: 0.6896 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.9639 - sparse_categorical_accuracy: 0.7010 - val_loss: 0.9174 - val_sparse_ca
tegorical_accuracy: 0.7176 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.9499 - sparse_categorical_accuracy: 0.7067 - val_loss: 0.8909 - val_sparse_ca
tegorical_accuracy: 0.7118 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.9654 - sparse_categorical_accuracy: 0.7037 - val_loss: 0.8671 - val_sparse_ca
tegorical_accuracy: 0.7029 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.9361 - sparse_categorical_accuracy: 0.7051 - val_loss: 0.9606 - val_sparse_ca
tegorical_accuracy: 0.6880 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.9361 - sparse_categorical_accuracy: 0.6986 - val_loss: 1.0948 - val_sparse_ca
tegorical_accuracy: 0.6439 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 0.9061 - sparse_categorical_accuracy: 0.7000 - val_loss: 0.9005 - val_sparse_ca
tegorical_accuracy: 0.6941 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.9183 - sparse_categorical_accuracy: 0.6984 - val_loss: 0.9350 - val_sparse_ca
tegorical_accuracy: 0.6966 - 1s/epoch - 2ms/step
Epoch 42/100
```

```
469/469 - 1s - loss: 0.9313 - sparse_categorical_accuracy: 0.6954 - val_loss: 0.9102 - val_sparse_ca
tegorical_accuracy: 0.6892 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 1.0145 - sparse_categorical_accuracy: 0.6822 - val_loss: 1.0944 - val_sparse_ca
tegorical_accuracy: 0.6936 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 1.0146 - sparse_categorical_accuracy: 0.6886 - val_loss: 0.9206 - val_sparse_ca
tegorical_accuracy: 0.6999 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.9473 - sparse_categorical_accuracy: 0.6960 - val_loss: 0.9287 - val_sparse_ca
tegorical_accuracy: 0.6994 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.9454 - sparse_categorical_accuracy: 0.6988 - val_loss: 0.9306 - val_sparse_ca
tegorical_accuracy: 0.6758 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 1.0176 - sparse_categorical_accuracy: 0.6970 - val_loss: 1.1976 - val_sparse_ca
tegorical_accuracy: 0.6871 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 1.0109 - sparse_categorical_accuracy: 0.6985 - val_loss: 1.0127 - val_sparse_ca
tegorical_accuracy: 0.6968 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 1.0459 - sparse_categorical_accuracy: 0.6936 - val_loss: 1.1644 - val_sparse_ca
tegorical_accuracy: 0.6858 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 1.0250 - sparse_categorical_accuracy: 0.6913 - val_loss: 1.0340 - val_sparse_ca
tegorical_accuracy: 0.6828 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.9774 - sparse_categorical_accuracy: 0.6962 - val_loss: 1.0176 - val_sparse_ca
tegorical_accuracy: 0.6927 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 1.0050 - sparse_categorical_accuracy: 0.6862 - val_loss: 0.9015 - val_sparse_ca
tegorical_accuracy: 0.7032 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.9865 - sparse_categorical_accuracy: 0.6917 - val_loss: 1.0859 - val_sparse_ca
tegorical_accuracy: 0.6846 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.9981 - sparse_categorical_accuracy: 0.6851 - val_loss: 0.9475 - val_sparse_ca
tegorical_accuracy: 0.6782 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 1.0121 - sparse_categorical_accuracy: 0.6885 - val_loss: 0.9800 - val_sparse_ca
tegorical_accuracy: 0.6718 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 1.0352 - sparse_categorical_accuracy: 0.6852 - val_loss: 1.5239 - val_sparse_ca
tegorical_accuracy: 0.4902 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 1.0588 - sparse_categorical_accuracy: 0.6780 - val_loss: 0.9739 - val_sparse_ca
tegorical_accuracy: 0.6747 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 1.0428 - sparse_categorical_accuracy: 0.6903 - val_loss: 1.0588 - val_sparse_ca
tegorical_accuracy: 0.6689 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 1.0841 - sparse_categorical_accuracy: 0.6678 - val_loss: 1.1056 - val_sparse_ca
tegorical_accuracy: 0.6518 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 1.0434 - sparse_categorical_accuracy: 0.6858 - val_loss: 1.1556 - val_sparse_ca
tegorical_accuracy: 0.6490 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 1.0703 - sparse_categorical_accuracy: 0.6907 - val_loss: 1.0199 - val_sparse_ca
tegorical_accuracy: 0.6614 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 1.0436 - sparse_categorical_accuracy: 0.6857 - val_loss: 1.0483 - val_sparse_ca
tegorical_accuracy: 0.6888 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 1.0855 - sparse_categorical_accuracy: 0.6883 - val_loss: 1.2118 - val_sparse_ca
tegorical_accuracy: 0.6945 - 1s/epoch - 3ms/step
Epoch 64/100
469/469 - 1s - loss: 1.1415 - sparse_categorical_accuracy: 0.6751 - val_loss: 1.3626 - val_sparse_ca
tegorical_accuracy: 0.5793 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 1.1107 - sparse_categorical_accuracy: 0.6624 - val_loss: 1.0752 - val_sparse_ca
tegorical_accuracy: 0.6820 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 1.0700 - sparse_categorical_accuracy: 0.6822 - val_loss: 1.3109 - val_sparse_ca
tegorical_accuracy: 0.6700 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 1.2776 - sparse_categorical_accuracy: 0.5700 - val_loss: 1.2158 - val_sparse_ca
tegorical_accuracy: 0.5644 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 1.3398 - sparse_categorical_accuracy: 0.5430 - val_loss: 1.3846 - val_sparse_ca
tegorical_accuracy: 0.4984 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 1.2580 - sparse_categorical_accuracy: 0.5634 - val_loss: 1.1170 - val_sparse_ca
tegorical_accuracy: 0.5931 - 1s/epoch - 2ms/step
```

```
Epoch 70/100
469/469 - 1s - loss: 1.2767 - sparse_categorical_accuracy: 0.6121 - val_loss: 1.1779 - val_sparse_ca
tegorical_accuracy: 0.6658 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 1.2016 - sparse_categorical_accuracy: 0.6643 - val_loss: 1.1901 - val_sparse_ca
tegorical_accuracy: 0.6964 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 1.0928 - sparse_categorical_accuracy: 0.6710 - val_loss: 1.0322 - val_sparse_ca
tegorical_accuracy: 0.6408 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 1.1237 - sparse_categorical_accuracy: 0.6685 - val_loss: 1.2371 - val_sparse_ca
tegorical_accuracy: 0.6434 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 1.1024 - sparse_categorical_accuracy: 0.6656 - val_loss: 1.1413 - val_sparse_ca
tegorical_accuracy: 0.6181 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 1.1552 - sparse_categorical_accuracy: 0.6726 - val_loss: 1.0502 - val_sparse_ca
tegorical_accuracy: 0.6770 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 1.1565 - sparse_categorical_accuracy: 0.6705 - val_loss: 1.1845 - val_sparse_ca
tegorical_accuracy: 0.6995 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 1.0956 - sparse_categorical_accuracy: 0.6690 - val_loss: 1.0604 - val_sparse_ca
tegorical_accuracy: 0.6847 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 1.0997 - sparse_categorical_accuracy: 0.6706 - val_loss: 1.1178 - val_sparse_ca
tegorical_accuracy: 0.6753 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 1.1939 - sparse_categorical_accuracy: 0.6796 - val_loss: 1.2386 - val_sparse_ca
tegorical_accuracy: 0.6733 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 1.2010 - sparse_categorical_accuracy: 0.6726 - val_loss: 1.3916 - val_sparse_ca
tegorical_accuracy: 0.6448 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 1.0727 - sparse_categorical_accuracy: 0.6731 - val_loss: 1.0811 - val_sparse_ca
tegorical_accuracy: 0.6513 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 1.1779 - sparse_categorical_accuracy: 0.6705 - val_loss: 1.0824 - val_sparse_ca
tegorical_accuracy: 0.6921 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 1.1526 - sparse_categorical_accuracy: 0.6566 - val_loss: 1.0545 - val_sparse_ca
tegorical_accuracy: 0.6471 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 1.0840 - sparse_categorical_accuracy: 0.6726 - val_loss: 1.0425 - val_sparse_ca
tegorical_accuracy: 0.6539 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 1.1478 - sparse_categorical_accuracy: 0.6822 - val_loss: 1.1040 - val_sparse_ca
tegorical_accuracy: 0.6415 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 1.1025 - sparse_categorical_accuracy: 0.6686 - val_loss: 1.0681 - val_sparse_ca
tegorical_accuracy: 0.6386 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 1.1034 - sparse_categorical_accuracy: 0.6698 - val_loss: 1.2824 - val_sparse_ca
tegorical_accuracy: 0.6644 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 1.2454 - sparse_categorical_accuracy: 0.6613 - val_loss: 1.2009 - val_sparse_ca
tegorical_accuracy: 0.6614 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 1.1428 - sparse_categorical_accuracy: 0.6616 - val_loss: 1.1210 - val_sparse_ca
tegorical_accuracy: 0.6353 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 1.1860 - sparse_categorical_accuracy: 0.6599 - val_loss: 1.1088 - val_sparse_ca
tegorical_accuracy: 0.6852 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 1.2232 - sparse_categorical_accuracy: 0.6592 - val_loss: 1.4925 - val_sparse_ca
tegorical_accuracy: 0.6546 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 1.2517 - sparse_categorical_accuracy: 0.6699 - val_loss: 1.5311 - val_sparse_ca
tegorical_accuracy: 0.6384 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 1.1903 - sparse_categorical_accuracy: 0.6630 - val_loss: 1.1443 - val_sparse_ca
tegorical_accuracy: 0.6662 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 1.1971 - sparse_categorical_accuracy: 0.6590 - val_loss: 1.6014 - val_sparse_ca
tegorical_accuracy: 0.6913 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 1.2104 - sparse_categorical_accuracy: 0.6616 - val_loss: 1.1777 - val_sparse_ca
tegorical_accuracy: 0.6408 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 1.1895 - sparse_categorical_accuracy: 0.6637 - val_loss: 1.3268 - val_sparse_ca
tegorical_accuracy: 0.6718 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 1.1468 - sparse_categorical_accuracy: 0.6603 - val_loss: 1.1299 - val_sparse_ca
```

```
tegorical_accuracy: 0.6315 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 1.2330 - sparse_categorical_accuracy: 0.6294 - val_loss: 1.1858 - val_sparse_ca
tegorical_accuracy: 0.5764 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 1.2194 - sparse_categorical_accuracy: 0.6124 - val_loss: 1.2830 - val_sparse_ca
tegorical_accuracy: 0.5934 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 1.1909 - sparse_categorical_accuracy: 0.6339 - val_loss: 1.1919 - val_sparse_ca
tegorical_accuracy: 0.6696 - 1s/epoch - 2ms/step
```

In [114]:

```
model6_f.evaluate(testX_f, testy_f, verbose=2)
```

```
313/313 - 0s - loss: 1.1919 - sparse_categorical_accuracy: 0.6696 - 392ms/epoch - 1ms/step
```

Out[114]:

```
[1.1918874979019165, 0.6696000099182129]
```

Model7: Regularization penalty is L1 with parameter 10^{-3} combined with L2 with parameter 10^{-5}

In [115]:

```
model7_f = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform',
  activity_regularizer = tf.keras.regularizers.L1L2(l1=0.001,l2=0.00001)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
# Fix the learning rate to be 0.001
model7_f.compile(optimizer=tf.keras.optimizers.Adam(0.001),
                 loss='sparse_categorical_crossentropy',
                 metrics=['sparse_categorical_accuracy']
                 )
history7_f = model7_f.fit(trainX_f, trainy_f,
                          batch_size=128,
                          epochs=100,
                          validation_data=(testX_f, testy_f),
                          verbose=2
                          )
```

```
Epoch 1/100
469/469 - 2s - loss: 2.0297 - sparse_categorical_accuracy: 0.1787 - val_loss: 1.8822 - val_sparse_ca
tegorical_accuracy: 0.2014 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 1.8181 - sparse_categorical_accuracy: 0.2395 - val_loss: 1.7216 - val_sparse_ca
tegorical_accuracy: 0.2767 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 1.6837 - sparse_categorical_accuracy: 0.3094 - val_loss: 1.6393 - val_sparse_ca
tegorical_accuracy: 0.3326 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 1.6342 - sparse_categorical_accuracy: 0.3313 - val_loss: 1.6031 - val_sparse_ca
tegorical_accuracy: 0.3555 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 1.6029 - sparse_categorical_accuracy: 0.3477 - val_loss: 1.5948 - val_sparse_ca
tegorical_accuracy: 0.3583 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 1.5442 - sparse_categorical_accuracy: 0.3813 - val_loss: 1.4383 - val_sparse_ca
tegorical_accuracy: 0.4338 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 1.4087 - sparse_categorical_accuracy: 0.4377 - val_loss: 1.8700 - val_sparse_ca
tegorical_accuracy: 0.3552 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 1.3794 - sparse_categorical_accuracy: 0.4422 - val_loss: 1.3506 - val_sparse_ca
tegorical_accuracy: 0.4494 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 1.3670 - sparse_categorical_accuracy: 0.4442 - val_loss: 1.3741 - val_sparse_ca
tegorical_accuracy: 0.4505 - 1s/epoch - 3ms/step
Epoch 10/100
469/469 - 1s - loss: 1.3609 - sparse_categorical_accuracy: 0.4448 - val_loss: 1.3697 - val_sparse_ca
tegorical_accuracy: 0.4390 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 1.3467 - sparse_categorical_accuracy: 0.4475 - val_loss: 1.3350 - val_sparse_ca
tegorical_accuracy: 0.4511 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 1.3440 - sparse_categorical_accuracy: 0.4480 - val_loss: 1.3502 - val_sparse_ca
tegorical_accuracy: 0.4503 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 1.3377 - sparse_categorical_accuracy: 0.4519 - val_loss: 1.3423 - val_sparse_ca
tegorical_accuracy: 0.4479 - 1s/epoch - 3ms/step
```

```
Epoch 14/100
469/469 - 1s - loss: 1.3382 - sparse_categorical_accuracy: 0.4527 - val_loss: 1.3993 - val_sparse_ca
tegorical_accuracy: 0.4504 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 1.3333 - sparse_categorical_accuracy: 0.4510 - val_loss: 1.3623 - val_sparse_ca
tegorical_accuracy: 0.4532 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 1.3279 - sparse_categorical_accuracy: 0.4546 - val_loss: 1.3203 - val_sparse_ca
tegorical_accuracy: 0.4508 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 1.3354 - sparse_categorical_accuracy: 0.4519 - val_loss: 1.3681 - val_sparse_ca
tegorical_accuracy: 0.4475 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 1.3160 - sparse_categorical_accuracy: 0.4568 - val_loss: 1.3378 - val_sparse_ca
tegorical_accuracy: 0.4507 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 1.3273 - sparse_categorical_accuracy: 0.4540 - val_loss: 1.3572 - val_sparse_ca
tegorical_accuracy: 0.4511 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 1.3315 - sparse_categorical_accuracy: 0.4524 - val_loss: 1.3347 - val_sparse_ca
tegorical_accuracy: 0.4505 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 1.3262 - sparse_categorical_accuracy: 0.4571 - val_loss: 1.3415 - val_sparse_ca
tegorical_accuracy: 0.4597 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 1.3333 - sparse_categorical_accuracy: 0.4505 - val_loss: 1.3259 - val_sparse_ca
tegorical_accuracy: 0.4569 - 1s/epoch - 3ms/step
Epoch 23/100
469/469 - 1s - loss: 1.3212 - sparse_categorical_accuracy: 0.4591 - val_loss: 1.3191 - val_sparse_ca
tegorical_accuracy: 0.4561 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 1.3224 - sparse_categorical_accuracy: 0.4560 - val_loss: 1.4042 - val_sparse_ca
tegorical_accuracy: 0.4647 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 1.3251 - sparse_categorical_accuracy: 0.4569 - val_loss: 1.3446 - val_sparse_ca
tegorical_accuracy: 0.4549 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 1.3309 - sparse_categorical_accuracy: 0.4510 - val_loss: 1.3635 - val_sparse_ca
tegorical_accuracy: 0.4533 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 1.3394 - sparse_categorical_accuracy: 0.4521 - val_loss: 1.3530 - val_sparse_ca
tegorical_accuracy: 0.4534 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 1.3151 - sparse_categorical_accuracy: 0.4599 - val_loss: 1.3164 - val_sparse_ca
tegorical_accuracy: 0.4575 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 1.3234 - sparse_categorical_accuracy: 0.4557 - val_loss: 1.3219 - val_sparse_ca
tegorical_accuracy: 0.4600 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 1.3249 - sparse_categorical_accuracy: 0.4552 - val_loss: 1.3439 - val_sparse_ca
tegorical_accuracy: 0.4587 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 1.3278 - sparse_categorical_accuracy: 0.4554 - val_loss: 1.3245 - val_sparse_ca
tegorical_accuracy: 0.4613 - 1s/epoch - 3ms/step
Epoch 32/100
469/469 - 1s - loss: 1.3215 - sparse_categorical_accuracy: 0.4581 - val_loss: 1.3280 - val_sparse_ca
tegorical_accuracy: 0.4610 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 1.3299 - sparse_categorical_accuracy: 0.4556 - val_loss: 1.5636 - val_sparse_ca
tegorical_accuracy: 0.4060 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 1.3358 - sparse_categorical_accuracy: 0.4546 - val_loss: 1.3418 - val_sparse_ca
tegorical_accuracy: 0.4523 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 1.3377 - sparse_categorical_accuracy: 0.4524 - val_loss: 1.3435 - val_sparse_ca
tegorical_accuracy: 0.4473 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 1.3351 - sparse_categorical_accuracy: 0.4541 - val_loss: 1.3248 - val_sparse_ca
tegorical_accuracy: 0.4589 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 1.3361 - sparse_categorical_accuracy: 0.4545 - val_loss: 1.3434 - val_sparse_ca
tegorical_accuracy: 0.4471 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 1.3232 - sparse_categorical_accuracy: 0.4573 - val_loss: 1.3502 - val_sparse_ca
tegorical_accuracy: 0.4484 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 1.3279 - sparse_categorical_accuracy: 0.4557 - val_loss: 1.3791 - val_sparse_ca
tegorical_accuracy: 0.4576 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 1.3499 - sparse_categorical_accuracy: 0.4512 - val_loss: 1.3360 - val_sparse_ca
tegorical_accuracy: 0.4541 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 1.3457 - sparse_categorical_accuracy: 0.4523 - val_loss: 1.3354 - val_sparse_ca
```

```
tegorical_accuracy: 0.4546 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 1.3364 - sparse_categorical_accuracy: 0.4543 - val_loss: 1.3422 - val_sparse_ca
tegorical_accuracy: 0.4529 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 1.3531 - sparse_categorical_accuracy: 0.4526 - val_loss: 1.3510 - val_sparse_ca
tegorical_accuracy: 0.4554 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 1.3429 - sparse_categorical_accuracy: 0.4535 - val_loss: 1.3571 - val_sparse_ca
tegorical_accuracy: 0.4528 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 1.3593 - sparse_categorical_accuracy: 0.4496 - val_loss: 1.3372 - val_sparse_ca
tegorical_accuracy: 0.4581 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 1.3778 - sparse_categorical_accuracy: 0.4496 - val_loss: 1.3951 - val_sparse_ca
tegorical_accuracy: 0.4367 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 1.3522 - sparse_categorical_accuracy: 0.4525 - val_loss: 1.3781 - val_sparse_ca
tegorical_accuracy: 0.4406 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 1.3513 - sparse_categorical_accuracy: 0.4525 - val_loss: 1.3503 - val_sparse_ca
tegorical_accuracy: 0.4527 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 1.3681 - sparse_categorical_accuracy: 0.4491 - val_loss: 1.3688 - val_sparse_ca
tegorical_accuracy: 0.4588 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 1.3713 - sparse_categorical_accuracy: 0.4484 - val_loss: 1.3386 - val_sparse_ca
tegorical_accuracy: 0.4523 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 1.3631 - sparse_categorical_accuracy: 0.4484 - val_loss: 1.5018 - val_sparse_ca
tegorical_accuracy: 0.4244 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 1.3494 - sparse_categorical_accuracy: 0.4531 - val_loss: 1.5274 - val_sparse_ca
tegorical_accuracy: 0.4411 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 1.3837 - sparse_categorical_accuracy: 0.4473 - val_loss: 1.3343 - val_sparse_ca
tegorical_accuracy: 0.4549 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 1.3631 - sparse_categorical_accuracy: 0.4508 - val_loss: 1.3783 - val_sparse_ca
tegorical_accuracy: 0.4470 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 1.3793 - sparse_categorical_accuracy: 0.4467 - val_loss: 1.4754 - val_sparse_ca
tegorical_accuracy: 0.4359 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 1.3661 - sparse_categorical_accuracy: 0.4487 - val_loss: 1.3393 - val_sparse_ca
tegorical_accuracy: 0.4567 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 1.3887 - sparse_categorical_accuracy: 0.4448 - val_loss: 1.4642 - val_sparse_ca
tegorical_accuracy: 0.4308 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 1.3922 - sparse_categorical_accuracy: 0.4451 - val_loss: 1.3855 - val_sparse_ca
tegorical_accuracy: 0.4456 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 1.3779 - sparse_categorical_accuracy: 0.4485 - val_loss: 1.3742 - val_sparse_ca
tegorical_accuracy: 0.4470 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 1.4000 - sparse_categorical_accuracy: 0.4438 - val_loss: 1.4609 - val_sparse_ca
tegorical_accuracy: 0.4119 - 1s/epoch - 3ms/step
Epoch 61/100
469/469 - 1s - loss: 1.4466 - sparse_categorical_accuracy: 0.4381 - val_loss: 1.8741 - val_sparse_ca
tegorical_accuracy: 0.2653 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 1.4661 - sparse_categorical_accuracy: 0.4199 - val_loss: 1.5075 - val_sparse_ca
tegorical_accuracy: 0.4098 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 1.3842 - sparse_categorical_accuracy: 0.4477 - val_loss: 1.3958 - val_sparse_ca
tegorical_accuracy: 0.4464 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 1.4166 - sparse_categorical_accuracy: 0.4427 - val_loss: 1.3754 - val_sparse_ca
tegorical_accuracy: 0.4482 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 1.4228 - sparse_categorical_accuracy: 0.4367 - val_loss: 1.3820 - val_sparse_ca
tegorical_accuracy: 0.4491 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 1.3844 - sparse_categorical_accuracy: 0.4474 - val_loss: 1.4319 - val_sparse_ca
tegorical_accuracy: 0.4456 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 1.4399 - sparse_categorical_accuracy: 0.4407 - val_loss: 1.5513 - val_sparse_ca
tegorical_accuracy: 0.3968 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 1.4225 - sparse_categorical_accuracy: 0.4390 - val_loss: 1.5065 - val_sparse_ca
tegorical_accuracy: 0.4147 - 1s/epoch - 3ms/step
Epoch 69/100
```

```
469/469 - 1s - loss: 1.4083 - sparse_categorical_accuracy: 0.4420 - val_loss: 1.4174 - val_sparse_ca
tegorical_accuracy: 0.4185 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 1.4104 - sparse_categorical_accuracy: 0.4397 - val_loss: 1.3886 - val_sparse_ca
tegorical_accuracy: 0.4484 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 1.4470 - sparse_categorical_accuracy: 0.4359 - val_loss: 1.4098 - val_sparse_ca
tegorical_accuracy: 0.4271 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 1.3977 - sparse_categorical_accuracy: 0.4424 - val_loss: 1.3949 - val_sparse_ca
tegorical_accuracy: 0.4469 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 1.4697 - sparse_categorical_accuracy: 0.4364 - val_loss: 1.3910 - val_sparse_ca
tegorical_accuracy: 0.4461 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 1.3829 - sparse_categorical_accuracy: 0.4468 - val_loss: 1.3854 - val_sparse_ca
tegorical_accuracy: 0.4519 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 1.4385 - sparse_categorical_accuracy: 0.4361 - val_loss: 1.6245 - val_sparse_ca
tegorical_accuracy: 0.4221 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 1.4167 - sparse_categorical_accuracy: 0.4419 - val_loss: 1.9781 - val_sparse_ca
tegorical_accuracy: 0.3849 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 1.4384 - sparse_categorical_accuracy: 0.4383 - val_loss: 1.4007 - val_sparse_ca
tegorical_accuracy: 0.4428 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 1.4599 - sparse_categorical_accuracy: 0.4334 - val_loss: 1.4888 - val_sparse_ca
tegorical_accuracy: 0.4123 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 1.5010 - sparse_categorical_accuracy: 0.4295 - val_loss: 1.3995 - val_sparse_ca
tegorical_accuracy: 0.4383 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 1.4461 - sparse_categorical_accuracy: 0.4349 - val_loss: 1.4277 - val_sparse_ca
tegorical_accuracy: 0.4345 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 1.4325 - sparse_categorical_accuracy: 0.4378 - val_loss: 1.3911 - val_sparse_ca
tegorical_accuracy: 0.4436 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 1.4463 - sparse_categorical_accuracy: 0.4339 - val_loss: 1.4032 - val_sparse_ca
tegorical_accuracy: 0.4495 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 1.4987 - sparse_categorical_accuracy: 0.4233 - val_loss: 1.4429 - val_sparse_ca
tegorical_accuracy: 0.4270 - 1s/epoch - 3ms/step
Epoch 84/100
469/469 - 1s - loss: 1.4506 - sparse_categorical_accuracy: 0.4309 - val_loss: 1.4423 - val_sparse_ca
tegorical_accuracy: 0.4472 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 1.4500 - sparse_categorical_accuracy: 0.4297 - val_loss: 1.4070 - val_sparse_ca
tegorical_accuracy: 0.4449 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 1.4456 - sparse_categorical_accuracy: 0.4368 - val_loss: 1.4569 - val_sparse_ca
tegorical_accuracy: 0.4416 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 1.5796 - sparse_categorical_accuracy: 0.4129 - val_loss: 1.4432 - val_sparse_ca
tegorical_accuracy: 0.4323 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 1.4828 - sparse_categorical_accuracy: 0.4230 - val_loss: 1.4904 - val_sparse_ca
tegorical_accuracy: 0.4048 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 1.5044 - sparse_categorical_accuracy: 0.4280 - val_loss: 1.4118 - val_sparse_ca
tegorical_accuracy: 0.4446 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 1.5166 - sparse_categorical_accuracy: 0.4158 - val_loss: 1.4837 - val_sparse_ca
tegorical_accuracy: 0.4133 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 1.5458 - sparse_categorical_accuracy: 0.4142 - val_loss: 1.5519 - val_sparse_ca
tegorical_accuracy: 0.4308 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 1.5135 - sparse_categorical_accuracy: 0.4253 - val_loss: 1.6239 - val_sparse_ca
tegorical_accuracy: 0.3874 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 1.5573 - sparse_categorical_accuracy: 0.4198 - val_loss: 1.4348 - val_sparse_ca
tegorical_accuracy: 0.4426 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 1.4791 - sparse_categorical_accuracy: 0.4364 - val_loss: 1.5274 - val_sparse_ca
tegorical_accuracy: 0.4327 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 1.4784 - sparse_categorical_accuracy: 0.4222 - val_loss: 1.6212 - val_sparse_ca
tegorical_accuracy: 0.4177 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 1.5430 - sparse_categorical_accuracy: 0.4120 - val_loss: 1.8271 - val_sparse_ca
tegorical_accuracy: 0.4296 - 1s/epoch - 2ms/step
```

```
Epoch 97/100
469/469 - 1s - loss: 1.5476 - sparse_categorical_accuracy: 0.4188 - val_loss: 1.6470 - val_sparse_ca
tegorical_accuracy: 0.3822 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 1.5099 - sparse_categorical_accuracy: 0.4237 - val_loss: 1.4244 - val_sparse_ca
tegorical_accuracy: 0.4420 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 1.5178 - sparse_categorical_accuracy: 0.4212 - val_loss: 1.5561 - val_sparse_ca
tegorical_accuracy: 0.4162 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 1.5812 - sparse_categorical_accuracy: 0.4047 - val_loss: 1.4799 - val_sparse_ca
tegorical_accuracy: 0.4182 - 1s/epoch - 2ms/step
```

In [117]:

```
model7_f.evaluate(testX_f, testy_f, verbose=2)
```

```
313/313 - 0s - loss: 1.4799 - sparse_categorical_accuracy: 0.4182 - 416ms/epoch - 1ms/step
```

Out[117]:

```
[1.4798961877822876, 0.41819998621940613]
```

Model8: Dropout with parameter 0.1

In [116]:

```
model8_f = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dropout(0.1),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
# Fix the learning rate to be 0.001
model8_f.compile(optimizer=tf.keras.optimizers.Adam(0.001),
               loss='sparse_categorical_crossentropy',
               metrics=['sparse_categorical_accuracy']
               )
history8_f = model8_f.fit(trainX_f, trainy_f,
                     batch_size=128,
                     epochs=100,
                     validation_data=(testX_f, testy_f),
                     verbose=2
                     )
```

```
Epoch 1/100
469/469 - 2s - loss: 0.6904 - sparse_categorical_accuracy: 0.7466 - val_loss: 0.5317 - val_sparse_ca
tegorical_accuracy: 0.8086 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 0.5110 - sparse_categorical_accuracy: 0.8140 - val_loss: 0.4644 - val_sparse_ca
tegorical_accuracy: 0.8277 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 0.4713 - sparse_categorical_accuracy: 0.8276 - val_loss: 0.4702 - val_sparse_ca
tegorical_accuracy: 0.8316 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 0.4564 - sparse_categorical_accuracy: 0.8305 - val_loss: 0.4417 - val_sparse_ca
tegorical_accuracy: 0.8380 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 0.4472 - sparse_categorical_accuracy: 0.8363 - val_loss: 0.4668 - val_sparse_ca
tegorical_accuracy: 0.8291 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 0.4360 - sparse_categorical_accuracy: 0.8390 - val_loss: 0.4404 - val_sparse_ca
tegorical_accuracy: 0.8387 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.4291 - sparse_categorical_accuracy: 0.8418 - val_loss: 0.4346 - val_sparse_ca
tegorical_accuracy: 0.8390 - 1s/epoch - 3ms/step
Epoch 8/100
469/469 - 1s - loss: 0.4231 - sparse_categorical_accuracy: 0.8443 - val_loss: 0.4322 - val_sparse_ca
tegorical_accuracy: 0.8451 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 0.4175 - sparse_categorical_accuracy: 0.8457 - val_loss: 0.4315 - val_sparse_ca
tegorical_accuracy: 0.8450 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 0.4146 - sparse_categorical_accuracy: 0.8468 - val_loss: 0.4251 - val_sparse_ca
tegorical_accuracy: 0.8463 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 0.4064 - sparse_categorical_accuracy: 0.8485 - val_loss: 0.4269 - val_sparse_ca
tegorical_accuracy: 0.8461 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.4088 - sparse_categorical_accuracy: 0.8479 - val_loss: 0.4190 - val_sparse_ca
tegorical_accuracy: 0.8457 - 1s/epoch - 2ms/step
Epoch 13/100
```

```
469/469 - 1s - loss: 0.4021 - sparse_categorical_accuracy: 0.8505 - val_loss: 0.4232 - val_sparse_ca
tegorical_accuracy: 0.8490 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.4032 - sparse_categorical_accuracy: 0.8507 - val_loss: 0.4203 - val_sparse_ca
tegorical_accuracy: 0.8479 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 0.4013 - sparse_categorical_accuracy: 0.8507 - val_loss: 0.4397 - val_sparse_ca
tegorical_accuracy: 0.8444 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.3979 - sparse_categorical_accuracy: 0.8527 - val_loss: 0.4355 - val_sparse_ca
tegorical_accuracy: 0.8402 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.3915 - sparse_categorical_accuracy: 0.8525 - val_loss: 0.4177 - val_sparse_ca
tegorical_accuracy: 0.8498 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 0.3948 - sparse_categorical_accuracy: 0.8528 - val_loss: 0.4588 - val_sparse_ca
tegorical_accuracy: 0.8418 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.3929 - sparse_categorical_accuracy: 0.8541 - val_loss: 0.4252 - val_sparse_ca
tegorical_accuracy: 0.8510 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.3895 - sparse_categorical_accuracy: 0.8543 - val_loss: 0.4160 - val_sparse_ca
tegorical_accuracy: 0.8495 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.3871 - sparse_categorical_accuracy: 0.8562 - val_loss: 0.4278 - val_sparse_ca
tegorical_accuracy: 0.8473 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.3841 - sparse_categorical_accuracy: 0.8556 - val_loss: 0.4152 - val_sparse_ca
tegorical_accuracy: 0.8536 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.3821 - sparse_categorical_accuracy: 0.8591 - val_loss: 0.4262 - val_sparse_ca
tegorical_accuracy: 0.8505 - 1s/epoch - 3ms/step
Epoch 24/100
469/469 - 1s - loss: 0.3810 - sparse_categorical_accuracy: 0.8570 - val_loss: 0.4234 - val_sparse_ca
tegorical_accuracy: 0.8504 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.3837 - sparse_categorical_accuracy: 0.8575 - val_loss: 0.4195 - val_sparse_ca
tegorical_accuracy: 0.8520 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.3795 - sparse_categorical_accuracy: 0.8588 - val_loss: 0.4295 - val_sparse_ca
tegorical_accuracy: 0.8467 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.3754 - sparse_categorical_accuracy: 0.8592 - val_loss: 0.4183 - val_sparse_ca
tegorical_accuracy: 0.8528 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.3762 - sparse_categorical_accuracy: 0.8614 - val_loss: 0.4060 - val_sparse_ca
tegorical_accuracy: 0.8551 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.3733 - sparse_categorical_accuracy: 0.8615 - val_loss: 0.4091 - val_sparse_ca
tegorical_accuracy: 0.8518 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.3704 - sparse_categorical_accuracy: 0.8608 - val_loss: 0.4236 - val_sparse_ca
tegorical_accuracy: 0.8472 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.3753 - sparse_categorical_accuracy: 0.8604 - val_loss: 0.4166 - val_sparse_ca
tegorical_accuracy: 0.8527 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.3719 - sparse_categorical_accuracy: 0.8600 - val_loss: 0.4279 - val_sparse_ca
tegorical_accuracy: 0.8468 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.3733 - sparse_categorical_accuracy: 0.8601 - val_loss: 0.4226 - val_sparse_ca
tegorical_accuracy: 0.8504 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.3729 - sparse_categorical_accuracy: 0.8604 - val_loss: 0.4225 - val_sparse_ca
tegorical_accuracy: 0.8490 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.3733 - sparse_categorical_accuracy: 0.8591 - val_loss: 0.4183 - val_sparse_ca
tegorical_accuracy: 0.8531 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.3685 - sparse_categorical_accuracy: 0.8625 - val_loss: 0.4264 - val_sparse_ca
tegorical_accuracy: 0.8517 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.3690 - sparse_categorical_accuracy: 0.8616 - val_loss: 0.4318 - val_sparse_ca
tegorical_accuracy: 0.8498 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.3691 - sparse_categorical_accuracy: 0.8627 - val_loss: 0.4245 - val_sparse_ca
tegorical_accuracy: 0.8507 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.3620 - sparse_categorical_accuracy: 0.8640 - val_loss: 0.4307 - val_sparse_ca
tegorical_accuracy: 0.8511 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 0.3682 - sparse_categorical_accuracy: 0.8617 - val_loss: 0.4279 - val_sparse_ca
tegorical_accuracy: 0.8491 - 1s/epoch - 2ms/step
```

```
Epoch 41/100
469/469 - 1s - loss: 0.3668 - sparse_categorical_accuracy: 0.8617 - val_loss: 0.4158 - val_sparse_ca
tegorical_accuracy: 0.8538 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.3653 - sparse_categorical_accuracy: 0.8630 - val_loss: 0.4329 - val_sparse_ca
tegorical_accuracy: 0.8510 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.3672 - sparse_categorical_accuracy: 0.8631 - val_loss: 0.4124 - val_sparse_ca
tegorical_accuracy: 0.8539 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.3632 - sparse_categorical_accuracy: 0.8642 - val_loss: 0.4383 - val_sparse_ca
tegorical_accuracy: 0.8452 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.3665 - sparse_categorical_accuracy: 0.8622 - val_loss: 0.4373 - val_sparse_ca
tegorical_accuracy: 0.8483 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.3625 - sparse_categorical_accuracy: 0.8657 - val_loss: 0.4278 - val_sparse_ca
tegorical_accuracy: 0.8493 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.3617 - sparse_categorical_accuracy: 0.8644 - val_loss: 0.4396 - val_sparse_ca
tegorical_accuracy: 0.8466 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.3621 - sparse_categorical_accuracy: 0.8645 - val_loss: 0.4471 - val_sparse_ca
tegorical_accuracy: 0.8431 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.3644 - sparse_categorical_accuracy: 0.8631 - val_loss: 0.4171 - val_sparse_ca
tegorical_accuracy: 0.8597 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.3581 - sparse_categorical_accuracy: 0.8667 - val_loss: 0.4153 - val_sparse_ca
tegorical_accuracy: 0.8544 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.3603 - sparse_categorical_accuracy: 0.8650 - val_loss: 0.4424 - val_sparse_ca
tegorical_accuracy: 0.8498 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.3583 - sparse_categorical_accuracy: 0.8657 - val_loss: 0.4290 - val_sparse_ca
tegorical_accuracy: 0.8533 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.3609 - sparse_categorical_accuracy: 0.8654 - val_loss: 0.4295 - val_sparse_ca
tegorical_accuracy: 0.8473 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.3561 - sparse_categorical_accuracy: 0.8667 - val_loss: 0.4415 - val_sparse_ca
tegorical_accuracy: 0.8505 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.3608 - sparse_categorical_accuracy: 0.8646 - val_loss: 0.4216 - val_sparse_ca
tegorical_accuracy: 0.8531 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.3561 - sparse_categorical_accuracy: 0.8661 - val_loss: 0.4306 - val_sparse_ca
tegorical_accuracy: 0.8457 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 0.3581 - sparse_categorical_accuracy: 0.8662 - val_loss: 0.4344 - val_sparse_ca
tegorical_accuracy: 0.8496 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.3567 - sparse_categorical_accuracy: 0.8656 - val_loss: 0.4234 - val_sparse_ca
tegorical_accuracy: 0.8530 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.3540 - sparse_categorical_accuracy: 0.8680 - val_loss: 0.4173 - val_sparse_ca
tegorical_accuracy: 0.8581 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.3591 - sparse_categorical_accuracy: 0.8657 - val_loss: 0.4222 - val_sparse_ca
tegorical_accuracy: 0.8523 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.3546 - sparse_categorical_accuracy: 0.8658 - val_loss: 0.4295 - val_sparse_ca
tegorical_accuracy: 0.8513 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.3541 - sparse_categorical_accuracy: 0.8668 - val_loss: 0.4288 - val_sparse_ca
tegorical_accuracy: 0.8540 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 0.3518 - sparse_categorical_accuracy: 0.8687 - val_loss: 0.4182 - val_sparse_ca
tegorical_accuracy: 0.8556 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.3548 - sparse_categorical_accuracy: 0.8667 - val_loss: 0.4355 - val_sparse_ca
tegorical_accuracy: 0.8465 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.3552 - sparse_categorical_accuracy: 0.8653 - val_loss: 0.4349 - val_sparse_ca
tegorical_accuracy: 0.8490 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.3552 - sparse_categorical_accuracy: 0.8653 - val_loss: 0.4544 - val_sparse_ca
tegorical_accuracy: 0.8435 - 1s/epoch - 3ms/step
Epoch 67/100
469/469 - 1s - loss: 0.3570 - sparse_categorical_accuracy: 0.8666 - val_loss: 0.4441 - val_sparse_ca
tegorical_accuracy: 0.8436 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 0.3504 - sparse_categorical_accuracy: 0.8694 - val_loss: 0.4337 - val_sparse_ca
```

```
tegorical_accuracy: 0.8524 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 0.3521 - sparse_categorical_accuracy: 0.8674 - val_loss: 0.4209 - val_sparse_ca
tegorical_accuracy: 0.8535 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.3510 - sparse_categorical_accuracy: 0.8681 - val_loss: 0.4296 - val_sparse_ca
tegorical_accuracy: 0.8493 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.3496 - sparse_categorical_accuracy: 0.8679 - val_loss: 0.4342 - val_sparse_ca
tegorical_accuracy: 0.8525 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.3525 - sparse_categorical_accuracy: 0.8673 - val_loss: 0.4282 - val_sparse_ca
tegorical_accuracy: 0.8532 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.3509 - sparse_categorical_accuracy: 0.8685 - val_loss: 0.4291 - val_sparse_ca
tegorical_accuracy: 0.8514 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.3540 - sparse_categorical_accuracy: 0.8674 - val_loss: 0.4268 - val_sparse_ca
tegorical_accuracy: 0.8497 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.3448 - sparse_categorical_accuracy: 0.8691 - val_loss: 0.4310 - val_sparse_ca
tegorical_accuracy: 0.8511 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.3514 - sparse_categorical_accuracy: 0.8687 - val_loss: 0.4281 - val_sparse_ca
tegorical_accuracy: 0.8544 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.3457 - sparse_categorical_accuracy: 0.8715 - val_loss: 0.4330 - val_sparse_ca
tegorical_accuracy: 0.8544 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.3486 - sparse_categorical_accuracy: 0.8698 - val_loss: 0.4286 - val_sparse_ca
tegorical_accuracy: 0.8546 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.3457 - sparse_categorical_accuracy: 0.8695 - val_loss: 0.4281 - val_sparse_ca
tegorical_accuracy: 0.8509 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.3498 - sparse_categorical_accuracy: 0.8698 - val_loss: 0.4184 - val_sparse_ca
tegorical_accuracy: 0.8555 - 1s/epoch - 3ms/step
Epoch 81/100
469/469 - 1s - loss: 0.3510 - sparse_categorical_accuracy: 0.8689 - val_loss: 0.4203 - val_sparse_ca
tegorical_accuracy: 0.8553 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.3473 - sparse_categorical_accuracy: 0.8709 - val_loss: 0.4161 - val_sparse_ca
tegorical_accuracy: 0.8580 - 1s/epoch - 3ms/step
Epoch 83/100
469/469 - 1s - loss: 0.3499 - sparse_categorical_accuracy: 0.8702 - val_loss: 0.4269 - val_sparse_ca
tegorical_accuracy: 0.8503 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.3468 - sparse_categorical_accuracy: 0.8711 - val_loss: 0.4419 - val_sparse_ca
tegorical_accuracy: 0.8481 - 1s/epoch - 3ms/step
Epoch 85/100
469/469 - 1s - loss: 0.3474 - sparse_categorical_accuracy: 0.8696 - val_loss: 0.4241 - val_sparse_ca
tegorical_accuracy: 0.8533 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.3439 - sparse_categorical_accuracy: 0.8714 - val_loss: 0.4479 - val_sparse_ca
tegorical_accuracy: 0.8487 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.3445 - sparse_categorical_accuracy: 0.8716 - val_loss: 0.4275 - val_sparse_ca
tegorical_accuracy: 0.8507 - 1s/epoch - 3ms/step
Epoch 88/100
469/469 - 1s - loss: 0.3456 - sparse_categorical_accuracy: 0.8694 - val_loss: 0.4552 - val_sparse_ca
tegorical_accuracy: 0.8458 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.3449 - sparse_categorical_accuracy: 0.8700 - val_loss: 0.4361 - val_sparse_ca
tegorical_accuracy: 0.8531 - 1s/epoch - 3ms/step
Epoch 90/100
469/469 - 1s - loss: 0.3475 - sparse_categorical_accuracy: 0.8696 - val_loss: 0.4428 - val_sparse_ca
tegorical_accuracy: 0.8514 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.3455 - sparse_categorical_accuracy: 0.8708 - val_loss: 0.4218 - val_sparse_ca
tegorical_accuracy: 0.8510 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.3419 - sparse_categorical_accuracy: 0.8720 - val_loss: 0.4385 - val_sparse_ca
tegorical_accuracy: 0.8474 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.3469 - sparse_categorical_accuracy: 0.8693 - val_loss: 0.4403 - val_sparse_ca
tegorical_accuracy: 0.8547 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.3458 - sparse_categorical_accuracy: 0.8708 - val_loss: 0.4343 - val_sparse_ca
tegorical_accuracy: 0.8498 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.3447 - sparse_categorical_accuracy: 0.8721 - val_loss: 0.4317 - val_sparse_ca
tegorical_accuracy: 0.8535 - 1s/epoch - 3ms/step
Epoch 96/100
```

```
469/469 - 1s - loss: 0.3413 - sparse_categorical_accuracy: 0.8723 - val_loss: 0.4268 - val_sparse_ca
tegorical_accuracy: 0.8571 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.3443 - sparse_categorical_accuracy: 0.8713 - val_loss: 0.4383 - val_sparse_ca
tegorical_accuracy: 0.8530 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.3461 - sparse_categorical_accuracy: 0.8709 - val_loss: 0.4261 - val_sparse_ca
tegorical_accuracy: 0.8544 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.3429 - sparse_categorical_accuracy: 0.8710 - val_loss: 0.4375 - val_sparse_ca
tegorical_accuracy: 0.8536 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.3476 - sparse_categorical_accuracy: 0.8695 - val_loss: 0.4405 - val_sparse_ca
tegorical_accuracy: 0.8517 - 1s/epoch - 3ms/step
```

In [118]:

```
model8_f.evaluate(testX_f, testy_f, verbose=2)
```

```
313/313 - 0s - loss: 0.4405 - sparse_categorical_accuracy: 0.8517 - 388ms/epoch - 1ms/step
```

Out[118]:

```
[0.44048720598220825, 0.8517000079154968]
```

Model9: Dropout with parameter 0.2

In [119]:

```
model9_f = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dropout(0.2),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
# Fix the learning rate to be 0.001
model9_f.compile(optimizer=tf.keras.optimizers.Adam(0.001),
                 loss='sparse_categorical_crossentropy',
                 metrics=['sparse_categorical_accuracy']
                 )
history9_f = model9_f.fit(trainX_f, trainy_f,
                      batch_size=128,
                      epochs=100,
                      validation_data=(testX_f, testy_f),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 2s - loss: 0.8805 - sparse_categorical_accuracy: 0.6604 - val_loss: 0.5771 - val_sparse_ca
tegorical_accuracy: 0.8000 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 0.6444 - sparse_categorical_accuracy: 0.7603 - val_loss: 0.5191 - val_sparse_ca
tegorical_accuracy: 0.8138 - 1s/epoch - 3ms/step
Epoch 3/100
469/469 - 1s - loss: 0.5999 - sparse_categorical_accuracy: 0.7768 - val_loss: 0.5381 - val_sparse_ca
tegorical_accuracy: 0.8184 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 0.5807 - sparse_categorical_accuracy: 0.7832 - val_loss: 0.4969 - val_sparse_ca
tegorical_accuracy: 0.8217 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 0.5690 - sparse_categorical_accuracy: 0.7883 - val_loss: 0.4771 - val_sparse_ca
tegorical_accuracy: 0.8334 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 0.5512 - sparse_categorical_accuracy: 0.7941 - val_loss: 0.4747 - val_sparse_ca
tegorical_accuracy: 0.8321 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.5457 - sparse_categorical_accuracy: 0.7957 - val_loss: 0.4540 - val_sparse_ca
tegorical_accuracy: 0.8411 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 0.5419 - sparse_categorical_accuracy: 0.7981 - val_loss: 0.4999 - val_sparse_ca
tegorical_accuracy: 0.8187 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 0.5340 - sparse_categorical_accuracy: 0.8005 - val_loss: 0.4599 - val_sparse_ca
tegorical_accuracy: 0.8332 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 0.5272 - sparse_categorical_accuracy: 0.8043 - val_loss: 0.4497 - val_sparse_ca
tegorical_accuracy: 0.8415 - 1s/epoch - 3ms/step
Epoch 11/100
469/469 - 1s - loss: 0.5187 - sparse_categorical_accuracy: 0.8067 - val_loss: 0.4551 - val_sparse_ca
tegorical_accuracy: 0.8378 - 1s/epoch - 2ms/step
Epoch 12/100
469/469 - 1s - loss: 0.5204 - sparse_categorical_accuracy: 0.8067 - val_loss: 0.4581 - val_sparse_ca
```

```
tegorical_accuracy: 0.8373 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.5129 - sparse_categorical_accuracy: 0.8088 - val_loss: 0.4468 - val_sparse_ca
tegorical_accuracy: 0.8396 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.5040 - sparse_categorical_accuracy: 0.8117 - val_loss: 0.4540 - val_sparse_ca
tegorical_accuracy: 0.8394 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 0.5087 - sparse_categorical_accuracy: 0.8100 - val_loss: 0.4440 - val_sparse_ca
tegorical_accuracy: 0.8407 - 1s/epoch - 2ms/step
Epoch 16/100
469/469 - 1s - loss: 0.5036 - sparse_categorical_accuracy: 0.8126 - val_loss: 0.4632 - val_sparse_ca
tegorical_accuracy: 0.8356 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.5002 - sparse_categorical_accuracy: 0.8125 - val_loss: 0.4594 - val_sparse_ca
tegorical_accuracy: 0.8378 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 0.4991 - sparse_categorical_accuracy: 0.8145 - val_loss: 0.4493 - val_sparse_ca
tegorical_accuracy: 0.8383 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.4956 - sparse_categorical_accuracy: 0.8159 - val_loss: 0.4584 - val_sparse_ca
tegorical_accuracy: 0.8375 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.4931 - sparse_categorical_accuracy: 0.8160 - val_loss: 0.4711 - val_sparse_ca
tegorical_accuracy: 0.8330 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.4937 - sparse_categorical_accuracy: 0.8159 - val_loss: 0.4753 - val_sparse_ca
tegorical_accuracy: 0.8268 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.4915 - sparse_categorical_accuracy: 0.8177 - val_loss: 0.4610 - val_sparse_ca
tegorical_accuracy: 0.8345 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.4942 - sparse_categorical_accuracy: 0.8173 - val_loss: 0.4593 - val_sparse_ca
tegorical_accuracy: 0.8370 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.4875 - sparse_categorical_accuracy: 0.8181 - val_loss: 0.4473 - val_sparse_ca
tegorical_accuracy: 0.8408 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.4832 - sparse_categorical_accuracy: 0.8212 - val_loss: 0.4553 - val_sparse_ca
tegorical_accuracy: 0.8391 - 1s/epoch - 2ms/step
Epoch 26/100
469/469 - 1s - loss: 0.4837 - sparse_categorical_accuracy: 0.8201 - val_loss: 0.4577 - val_sparse_ca
tegorical_accuracy: 0.8409 - 1s/epoch - 2ms/step
Epoch 27/100
469/469 - 1s - loss: 0.4816 - sparse_categorical_accuracy: 0.8229 - val_loss: 0.4382 - val_sparse_ca
tegorical_accuracy: 0.8439 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.4830 - sparse_categorical_accuracy: 0.8206 - val_loss: 0.4661 - val_sparse_ca
tegorical_accuracy: 0.8310 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.4790 - sparse_categorical_accuracy: 0.8227 - val_loss: 0.4633 - val_sparse_ca
tegorical_accuracy: 0.8331 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.4794 - sparse_categorical_accuracy: 0.8222 - val_loss: 0.4425 - val_sparse_ca
tegorical_accuracy: 0.8456 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.4742 - sparse_categorical_accuracy: 0.8255 - val_loss: 0.4556 - val_sparse_ca
tegorical_accuracy: 0.8381 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.4781 - sparse_categorical_accuracy: 0.8228 - val_loss: 0.4493 - val_sparse_ca
tegorical_accuracy: 0.8395 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.4783 - sparse_categorical_accuracy: 0.8237 - val_loss: 0.4468 - val_sparse_ca
tegorical_accuracy: 0.8443 - 1s/epoch - 2ms/step
Epoch 34/100
469/469 - 1s - loss: 0.4759 - sparse_categorical_accuracy: 0.8233 - val_loss: 0.4661 - val_sparse_ca
tegorical_accuracy: 0.8346 - 1s/epoch - 2ms/step
Epoch 35/100
469/469 - 1s - loss: 0.4717 - sparse_categorical_accuracy: 0.8249 - val_loss: 0.4657 - val_sparse_ca
tegorical_accuracy: 0.8314 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.4770 - sparse_categorical_accuracy: 0.8227 - val_loss: 0.4709 - val_sparse_ca
tegorical_accuracy: 0.8335 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.4690 - sparse_categorical_accuracy: 0.8256 - val_loss: 0.4378 - val_sparse_ca
tegorical_accuracy: 0.8432 - 1s/epoch - 3ms/step
Epoch 38/100
469/469 - 1s - loss: 0.4719 - sparse_categorical_accuracy: 0.8255 - val_loss: 0.4656 - val_sparse_ca
tegorical_accuracy: 0.8280 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.4745 - sparse_categorical_accuracy: 0.8252 - val_loss: 0.4546 - val_sparse_ca
tegorical_accuracy: 0.8402 - 1s/epoch - 2ms/step
Epoch 40/100
```

```
469/469 - 1s - loss: 0.4687 - sparse_categorical_accuracy: 0.8251 - val_loss: 0.4501 - val_sparse_ca
tegorical_accuracy: 0.8439 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.4720 - sparse_categorical_accuracy: 0.8270 - val_loss: 0.4503 - val_sparse_ca
tegorical_accuracy: 0.8421 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.4703 - sparse_categorical_accuracy: 0.8273 - val_loss: 0.4642 - val_sparse_ca
tegorical_accuracy: 0.8386 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.4687 - sparse_categorical_accuracy: 0.8281 - val_loss: 0.4417 - val_sparse_ca
tegorical_accuracy: 0.8423 - 1s/epoch - 3ms/step
Epoch 44/100
469/469 - 1s - loss: 0.4632 - sparse_categorical_accuracy: 0.8305 - val_loss: 0.4616 - val_sparse_ca
tegorical_accuracy: 0.8395 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.4650 - sparse_categorical_accuracy: 0.8278 - val_loss: 0.4418 - val_sparse_ca
tegorical_accuracy: 0.8407 - 1s/epoch - 3ms/step
Epoch 46/100
469/469 - 1s - loss: 0.4661 - sparse_categorical_accuracy: 0.8287 - val_loss: 0.4452 - val_sparse_ca
tegorical_accuracy: 0.8443 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.4619 - sparse_categorical_accuracy: 0.8302 - val_loss: 0.4555 - val_sparse_ca
tegorical_accuracy: 0.8389 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.4662 - sparse_categorical_accuracy: 0.8278 - val_loss: 0.4640 - val_sparse_ca
tegorical_accuracy: 0.8334 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.4654 - sparse_categorical_accuracy: 0.8286 - val_loss: 0.4579 - val_sparse_ca
tegorical_accuracy: 0.8410 - 1s/epoch - 3ms/step
Epoch 50/100
469/469 - 1s - loss: 0.4633 - sparse_categorical_accuracy: 0.8286 - val_loss: 0.4526 - val_sparse_ca
tegorical_accuracy: 0.8417 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.4627 - sparse_categorical_accuracy: 0.8323 - val_loss: 0.4730 - val_sparse_ca
tegorical_accuracy: 0.8420 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.4581 - sparse_categorical_accuracy: 0.8306 - val_loss: 0.4415 - val_sparse_ca
tegorical_accuracy: 0.8424 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.4672 - sparse_categorical_accuracy: 0.8286 - val_loss: 0.4507 - val_sparse_ca
tegorical_accuracy: 0.8398 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.4639 - sparse_categorical_accuracy: 0.8295 - val_loss: 0.4555 - val_sparse_ca
tegorical_accuracy: 0.8432 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.4637 - sparse_categorical_accuracy: 0.8316 - val_loss: 0.4469 - val_sparse_ca
tegorical_accuracy: 0.8421 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.4649 - sparse_categorical_accuracy: 0.8295 - val_loss: 0.4571 - val_sparse_ca
tegorical_accuracy: 0.8352 - 1s/epoch - 2ms/step
Epoch 57/100
469/469 - 1s - loss: 0.4623 - sparse_categorical_accuracy: 0.8303 - val_loss: 0.4715 - val_sparse_ca
tegorical_accuracy: 0.8304 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.4648 - sparse_categorical_accuracy: 0.8301 - val_loss: 0.4454 - val_sparse_ca
tegorical_accuracy: 0.8418 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.4604 - sparse_categorical_accuracy: 0.8304 - val_loss: 0.4472 - val_sparse_ca
tegorical_accuracy: 0.8371 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.4619 - sparse_categorical_accuracy: 0.8295 - val_loss: 0.4615 - val_sparse_ca
tegorical_accuracy: 0.8319 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.4621 - sparse_categorical_accuracy: 0.8311 - val_loss: 0.4706 - val_sparse_ca
tegorical_accuracy: 0.8267 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.4609 - sparse_categorical_accuracy: 0.8313 - val_loss: 0.4479 - val_sparse_ca
tegorical_accuracy: 0.8455 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 0.4585 - sparse_categorical_accuracy: 0.8315 - val_loss: 0.4564 - val_sparse_ca
tegorical_accuracy: 0.8394 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.4586 - sparse_categorical_accuracy: 0.8318 - val_loss: 0.4535 - val_sparse_ca
tegorical_accuracy: 0.8424 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.4610 - sparse_categorical_accuracy: 0.8305 - val_loss: 0.4668 - val_sparse_ca
tegorical_accuracy: 0.8330 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.4565 - sparse_categorical_accuracy: 0.8313 - val_loss: 0.4491 - val_sparse_ca
tegorical_accuracy: 0.8430 - 1s/epoch - 2ms/step
Epoch 67/100
469/469 - 1s - loss: 0.4542 - sparse_categorical_accuracy: 0.8319 - val_loss: 0.4598 - val_sparse_ca
tegorical_accuracy: 0.8349 - 1s/epoch - 3ms/step
```

```
Epoch 68/100
469/469 - 1s - loss: 0.4615 - sparse_categorical_accuracy: 0.8322 - val_loss: 0.4612 - val_sparse_ca
tegorical_accuracy: 0.8372 - 1s/epoch - 2ms/step
Epoch 69/100
469/469 - 1s - loss: 0.4604 - sparse_categorical_accuracy: 0.8317 - val_loss: 0.4507 - val_sparse_ca
tegorical_accuracy: 0.8401 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.4595 - sparse_categorical_accuracy: 0.8326 - val_loss: 0.4648 - val_sparse_ca
tegorical_accuracy: 0.8299 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.4643 - sparse_categorical_accuracy: 0.8305 - val_loss: 0.4640 - val_sparse_ca
tegorical_accuracy: 0.8366 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.4568 - sparse_categorical_accuracy: 0.8328 - val_loss: 0.4522 - val_sparse_ca
tegorical_accuracy: 0.8422 - 1s/epoch - 2ms/step
Epoch 73/100
469/469 - 1s - loss: 0.4564 - sparse_categorical_accuracy: 0.8350 - val_loss: 0.4641 - val_sparse_ca
tegorical_accuracy: 0.8395 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.4559 - sparse_categorical_accuracy: 0.8327 - val_loss: 0.4551 - val_sparse_ca
tegorical_accuracy: 0.8400 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.4682 - sparse_categorical_accuracy: 0.8273 - val_loss: 0.5041 - val_sparse_ca
tegorical_accuracy: 0.8304 - 1s/epoch - 3ms/step
Epoch 76/100
469/469 - 1s - loss: 0.4590 - sparse_categorical_accuracy: 0.8318 - val_loss: 0.4571 - val_sparse_ca
tegorical_accuracy: 0.8421 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.4551 - sparse_categorical_accuracy: 0.8328 - val_loss: 0.4645 - val_sparse_ca
tegorical_accuracy: 0.8401 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.4596 - sparse_categorical_accuracy: 0.8308 - val_loss: 0.4416 - val_sparse_ca
tegorical_accuracy: 0.8425 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.4544 - sparse_categorical_accuracy: 0.8354 - val_loss: 0.4685 - val_sparse_ca
tegorical_accuracy: 0.8375 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.4570 - sparse_categorical_accuracy: 0.8322 - val_loss: 0.4651 - val_sparse_ca
tegorical_accuracy: 0.8363 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.4505 - sparse_categorical_accuracy: 0.8341 - val_loss: 0.4446 - val_sparse_ca
tegorical_accuracy: 0.8462 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.4544 - sparse_categorical_accuracy: 0.8345 - val_loss: 0.4685 - val_sparse_ca
tegorical_accuracy: 0.8307 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.4545 - sparse_categorical_accuracy: 0.8323 - val_loss: 0.4603 - val_sparse_ca
tegorical_accuracy: 0.8420 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.4551 - sparse_categorical_accuracy: 0.8326 - val_loss: 0.4506 - val_sparse_ca
tegorical_accuracy: 0.8406 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.4550 - sparse_categorical_accuracy: 0.8338 - val_loss: 0.4412 - val_sparse_ca
tegorical_accuracy: 0.8470 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.4537 - sparse_categorical_accuracy: 0.8342 - val_loss: 0.4919 - val_sparse_ca
tegorical_accuracy: 0.8185 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.4509 - sparse_categorical_accuracy: 0.8351 - val_loss: 0.4637 - val_sparse_ca
tegorical_accuracy: 0.8346 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.4516 - sparse_categorical_accuracy: 0.8342 - val_loss: 0.4442 - val_sparse_ca
tegorical_accuracy: 0.8459 - 1s/epoch - 2ms/step
Epoch 89/100
469/469 - 1s - loss: 0.4580 - sparse_categorical_accuracy: 0.8317 - val_loss: 0.4631 - val_sparse_ca
tegorical_accuracy: 0.8430 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.4573 - sparse_categorical_accuracy: 0.8324 - val_loss: 0.4457 - val_sparse_ca
tegorical_accuracy: 0.8473 - 1s/epoch - 2ms/step
Epoch 91/100
469/469 - 1s - loss: 0.4504 - sparse_categorical_accuracy: 0.8336 - val_loss: 0.4571 - val_sparse_ca
tegorical_accuracy: 0.8415 - 1s/epoch - 2ms/step
Epoch 92/100
469/469 - 1s - loss: 0.4554 - sparse_categorical_accuracy: 0.8324 - val_loss: 0.4785 - val_sparse_ca
tegorical_accuracy: 0.8319 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.4531 - sparse_categorical_accuracy: 0.8339 - val_loss: 0.4468 - val_sparse_ca
tegorical_accuracy: 0.8433 - 1s/epoch - 2ms/step
Epoch 94/100
469/469 - 1s - loss: 0.4537 - sparse_categorical_accuracy: 0.8339 - val_loss: 0.4607 - val_sparse_ca
tegorical_accuracy: 0.8361 - 1s/epoch - 2ms/step
Epoch 95/100
469/469 - 1s - loss: 0.4559 - sparse_categorical_accuracy: 0.8315 - val_loss: 0.4784 - val_sparse_ca
```

```
tegorical_accuracy: 0.8339 - 1s/epoch - 2ms/step
Epoch 96/100
469/469 - 1s - loss: 0.4527 - sparse_categorical_accuracy: 0.8332 - val_loss: 0.4597 - val_sparse_ca
tegorical_accuracy: 0.8373 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.4521 - sparse_categorical_accuracy: 0.8343 - val_loss: 0.4465 - val_sparse_ca
tegorical_accuracy: 0.8418 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.4526 - sparse_categorical_accuracy: 0.8339 - val_loss: 0.4684 - val_sparse_ca
tegorical_accuracy: 0.8312 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.4548 - sparse_categorical_accuracy: 0.8322 - val_loss: 0.4456 - val_sparse_ca
tegorical_accuracy: 0.8428 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.4499 - sparse_categorical_accuracy: 0.8356 - val_loss: 0.4757 - val_sparse_ca
tegorical_accuracy: 0.8375 - 1s/epoch - 2ms/step
```

In [120]:

```
model9_f.evaluate(testX_f, testy_f, verbose=2)
```

```
313/313 - 0s - loss: 0.4757 - sparse_categorical_accuracy: 0.8375 - 378ms/epoch - 1ms/step
```

Out[120]:

```
[0.47566211223602295, 0.8374999761581421]
```

Model10: Dropout with parameter 0.3

In [121]:

```
model10_f = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dropout(0.3),
  tf.keras.layers.Dense(16, activation='relu', kernel_initializer='random_uniform'),
  tf.keras.layers.Dense(10, activation='softmax', kernel_initializer='random_uniform')
])
# Fix the learning rate to be 0.001
model10_f.compile(optimizer=tf.keras.optimizers.Adam(0.001),
                loss='sparse_categorical_crossentropy',
                metrics=['sparse_categorical_accuracy']
                )
history10_f = model10_f.fit(trainX_f, trainy_f,
                      batch_size=128,
                      epochs=100,
                      validation_data=(testX_f, testy_f),
                      verbose=2
                      )
```

```
Epoch 1/100
469/469 - 2s - loss: 1.0137 - sparse_categorical_accuracy: 0.5928 - val_loss: 0.6270 - val_sparse_ca
tegorical_accuracy: 0.7804 - 2s/epoch - 4ms/step
Epoch 2/100
469/469 - 1s - loss: 0.7832 - sparse_categorical_accuracy: 0.6894 - val_loss: 0.5934 - val_sparse_ca
tegorical_accuracy: 0.8042 - 1s/epoch - 2ms/step
Epoch 3/100
469/469 - 1s - loss: 0.7379 - sparse_categorical_accuracy: 0.7122 - val_loss: 0.5791 - val_sparse_ca
tegorical_accuracy: 0.7931 - 1s/epoch - 2ms/step
Epoch 4/100
469/469 - 1s - loss: 0.7105 - sparse_categorical_accuracy: 0.7266 - val_loss: 0.5494 - val_sparse_ca
tegorical_accuracy: 0.8166 - 1s/epoch - 2ms/step
Epoch 5/100
469/469 - 1s - loss: 0.6899 - sparse_categorical_accuracy: 0.7371 - val_loss: 0.5580 - val_sparse_ca
tegorical_accuracy: 0.8094 - 1s/epoch - 2ms/step
Epoch 6/100
469/469 - 1s - loss: 0.6777 - sparse_categorical_accuracy: 0.7422 - val_loss: 0.5251 - val_sparse_ca
tegorical_accuracy: 0.8183 - 1s/epoch - 2ms/step
Epoch 7/100
469/469 - 1s - loss: 0.6659 - sparse_categorical_accuracy: 0.7465 - val_loss: 0.5497 - val_sparse_ca
tegorical_accuracy: 0.8121 - 1s/epoch - 2ms/step
Epoch 8/100
469/469 - 1s - loss: 0.6576 - sparse_categorical_accuracy: 0.7508 - val_loss: 0.5436 - val_sparse_ca
tegorical_accuracy: 0.8182 - 1s/epoch - 2ms/step
Epoch 9/100
469/469 - 1s - loss: 0.6519 - sparse_categorical_accuracy: 0.7545 - val_loss: 0.5485 - val_sparse_ca
tegorical_accuracy: 0.8101 - 1s/epoch - 2ms/step
Epoch 10/100
469/469 - 1s - loss: 0.6417 - sparse_categorical_accuracy: 0.7582 - val_loss: 0.5112 - val_sparse_ca
tegorical_accuracy: 0.8211 - 1s/epoch - 2ms/step
Epoch 11/100
469/469 - 1s - loss: 0.6458 - sparse_categorical_accuracy: 0.7585 - val_loss: 0.5699 - val_sparse_ca
tegorical_accuracy: 0.8030 - 1s/epoch - 2ms/step
```

```
Epoch 12/100
469/469 - 1s - loss: 0.6425 - sparse_categorical_accuracy: 0.7586 - val_loss: 0.5677 - val_sparse_ca
tegorical_accuracy: 0.8140 - 1s/epoch - 2ms/step
Epoch 13/100
469/469 - 1s - loss: 0.6398 - sparse_categorical_accuracy: 0.7596 - val_loss: 0.5249 - val_sparse_ca
tegorical_accuracy: 0.8178 - 1s/epoch - 2ms/step
Epoch 14/100
469/469 - 1s - loss: 0.6300 - sparse_categorical_accuracy: 0.7664 - val_loss: 0.5376 - val_sparse_ca
tegorical_accuracy: 0.8048 - 1s/epoch - 2ms/step
Epoch 15/100
469/469 - 1s - loss: 0.6295 - sparse_categorical_accuracy: 0.7674 - val_loss: 0.5231 - val_sparse_ca
tegorical_accuracy: 0.8237 - 1s/epoch - 3ms/step
Epoch 16/100
469/469 - 1s - loss: 0.6322 - sparse_categorical_accuracy: 0.7652 - val_loss: 0.5258 - val_sparse_ca
tegorical_accuracy: 0.8163 - 1s/epoch - 2ms/step
Epoch 17/100
469/469 - 1s - loss: 0.6275 - sparse_categorical_accuracy: 0.7670 - val_loss: 0.5655 - val_sparse_ca
tegorical_accuracy: 0.7957 - 1s/epoch - 2ms/step
Epoch 18/100
469/469 - 1s - loss: 0.6229 - sparse_categorical_accuracy: 0.7698 - val_loss: 0.5473 - val_sparse_ca
tegorical_accuracy: 0.8093 - 1s/epoch - 2ms/step
Epoch 19/100
469/469 - 1s - loss: 0.6152 - sparse_categorical_accuracy: 0.7704 - val_loss: 0.5455 - val_sparse_ca
tegorical_accuracy: 0.8101 - 1s/epoch - 2ms/step
Epoch 20/100
469/469 - 1s - loss: 0.6157 - sparse_categorical_accuracy: 0.7714 - val_loss: 0.5034 - val_sparse_ca
tegorical_accuracy: 0.8241 - 1s/epoch - 2ms/step
Epoch 21/100
469/469 - 1s - loss: 0.6132 - sparse_categorical_accuracy: 0.7700 - val_loss: 0.5152 - val_sparse_ca
tegorical_accuracy: 0.8234 - 1s/epoch - 2ms/step
Epoch 22/100
469/469 - 1s - loss: 0.6111 - sparse_categorical_accuracy: 0.7736 - val_loss: 0.5084 - val_sparse_ca
tegorical_accuracy: 0.8227 - 1s/epoch - 2ms/step
Epoch 23/100
469/469 - 1s - loss: 0.6135 - sparse_categorical_accuracy: 0.7711 - val_loss: 0.5236 - val_sparse_ca
tegorical_accuracy: 0.8210 - 1s/epoch - 2ms/step
Epoch 24/100
469/469 - 1s - loss: 0.6108 - sparse_categorical_accuracy: 0.7769 - val_loss: 0.5089 - val_sparse_ca
tegorical_accuracy: 0.8272 - 1s/epoch - 2ms/step
Epoch 25/100
469/469 - 1s - loss: 0.6126 - sparse_categorical_accuracy: 0.7747 - val_loss: 0.5261 - val_sparse_ca
tegorical_accuracy: 0.8242 - 1s/epoch - 3ms/step
Epoch 26/100
469/469 - 1s - loss: 0.6113 - sparse_categorical_accuracy: 0.7750 - val_loss: 0.4916 - val_sparse_ca
tegorical_accuracy: 0.8305 - 1s/epoch - 3ms/step
Epoch 27/100
469/469 - 1s - loss: 0.6069 - sparse_categorical_accuracy: 0.7758 - val_loss: 0.5273 - val_sparse_ca
tegorical_accuracy: 0.8263 - 1s/epoch - 2ms/step
Epoch 28/100
469/469 - 1s - loss: 0.6071 - sparse_categorical_accuracy: 0.7752 - val_loss: 0.4990 - val_sparse_ca
tegorical_accuracy: 0.8281 - 1s/epoch - 2ms/step
Epoch 29/100
469/469 - 1s - loss: 0.6005 - sparse_categorical_accuracy: 0.7764 - val_loss: 0.5465 - val_sparse_ca
tegorical_accuracy: 0.8132 - 1s/epoch - 2ms/step
Epoch 30/100
469/469 - 1s - loss: 0.6027 - sparse_categorical_accuracy: 0.7760 - val_loss: 0.5313 - val_sparse_ca
tegorical_accuracy: 0.8161 - 1s/epoch - 2ms/step
Epoch 31/100
469/469 - 1s - loss: 0.6013 - sparse_categorical_accuracy: 0.7767 - val_loss: 0.5249 - val_sparse_ca
tegorical_accuracy: 0.8190 - 1s/epoch - 2ms/step
Epoch 32/100
469/469 - 1s - loss: 0.6090 - sparse_categorical_accuracy: 0.7744 - val_loss: 0.5158 - val_sparse_ca
tegorical_accuracy: 0.8208 - 1s/epoch - 2ms/step
Epoch 33/100
469/469 - 1s - loss: 0.6087 - sparse_categorical_accuracy: 0.7746 - val_loss: 0.5584 - val_sparse_ca
tegorical_accuracy: 0.8012 - 1s/epoch - 3ms/step
Epoch 34/100
469/469 - 1s - loss: 0.6003 - sparse_categorical_accuracy: 0.7796 - val_loss: 0.5217 - val_sparse_ca
tegorical_accuracy: 0.8235 - 1s/epoch - 3ms/step
Epoch 35/100
469/469 - 1s - loss: 0.6003 - sparse_categorical_accuracy: 0.7803 - val_loss: 0.5122 - val_sparse_ca
tegorical_accuracy: 0.8257 - 1s/epoch - 2ms/step
Epoch 36/100
469/469 - 1s - loss: 0.5994 - sparse_categorical_accuracy: 0.7785 - val_loss: 0.5100 - val_sparse_ca
tegorical_accuracy: 0.8215 - 1s/epoch - 2ms/step
Epoch 37/100
469/469 - 1s - loss: 0.6001 - sparse_categorical_accuracy: 0.7785 - val_loss: 0.5323 - val_sparse_ca
tegorical_accuracy: 0.8161 - 1s/epoch - 2ms/step
Epoch 38/100
469/469 - 1s - loss: 0.6032 - sparse_categorical_accuracy: 0.7760 - val_loss: 0.5056 - val_sparse_ca
tegorical_accuracy: 0.8281 - 1s/epoch - 2ms/step
Epoch 39/100
469/469 - 1s - loss: 0.5971 - sparse_categorical_accuracy: 0.7803 - val_loss: 0.5350 - val_sparse_ca
```

```
tegorical_accuracy: 0.8115 - 1s/epoch - 2ms/step
Epoch 40/100
469/469 - 1s - loss: 0.5976 - sparse_categorical_accuracy: 0.7801 - val_loss: 0.5327 - val_sparse_ca
tegorical_accuracy: 0.8202 - 1s/epoch - 2ms/step
Epoch 41/100
469/469 - 1s - loss: 0.6005 - sparse_categorical_accuracy: 0.7793 - val_loss: 0.5698 - val_sparse_ca
tegorical_accuracy: 0.7995 - 1s/epoch - 2ms/step
Epoch 42/100
469/469 - 1s - loss: 0.5910 - sparse_categorical_accuracy: 0.7813 - val_loss: 0.5209 - val_sparse_ca
tegorical_accuracy: 0.8184 - 1s/epoch - 2ms/step
Epoch 43/100
469/469 - 1s - loss: 0.5940 - sparse_categorical_accuracy: 0.7808 - val_loss: 0.5756 - val_sparse_ca
tegorical_accuracy: 0.7979 - 1s/epoch - 2ms/step
Epoch 44/100
469/469 - 1s - loss: 0.5997 - sparse_categorical_accuracy: 0.7780 - val_loss: 0.5087 - val_sparse_ca
tegorical_accuracy: 0.8226 - 1s/epoch - 2ms/step
Epoch 45/100
469/469 - 1s - loss: 0.5878 - sparse_categorical_accuracy: 0.7836 - val_loss: 0.5138 - val_sparse_ca
tegorical_accuracy: 0.8216 - 1s/epoch - 2ms/step
Epoch 46/100
469/469 - 1s - loss: 0.5936 - sparse_categorical_accuracy: 0.7807 - val_loss: 0.5839 - val_sparse_ca
tegorical_accuracy: 0.8012 - 1s/epoch - 2ms/step
Epoch 47/100
469/469 - 1s - loss: 0.5925 - sparse_categorical_accuracy: 0.7814 - val_loss: 0.5245 - val_sparse_ca
tegorical_accuracy: 0.8188 - 1s/epoch - 2ms/step
Epoch 48/100
469/469 - 1s - loss: 0.5828 - sparse_categorical_accuracy: 0.7853 - val_loss: 0.5408 - val_sparse_ca
tegorical_accuracy: 0.8175 - 1s/epoch - 2ms/step
Epoch 49/100
469/469 - 1s - loss: 0.5995 - sparse_categorical_accuracy: 0.7804 - val_loss: 0.5145 - val_sparse_ca
tegorical_accuracy: 0.8279 - 1s/epoch - 2ms/step
Epoch 50/100
469/469 - 1s - loss: 0.5919 - sparse_categorical_accuracy: 0.7833 - val_loss: 0.5206 - val_sparse_ca
tegorical_accuracy: 0.8235 - 1s/epoch - 2ms/step
Epoch 51/100
469/469 - 1s - loss: 0.5845 - sparse_categorical_accuracy: 0.7828 - val_loss: 0.5204 - val_sparse_ca
tegorical_accuracy: 0.8224 - 1s/epoch - 2ms/step
Epoch 52/100
469/469 - 1s - loss: 0.5931 - sparse_categorical_accuracy: 0.7821 - val_loss: 0.5331 - val_sparse_ca
tegorical_accuracy: 0.8208 - 1s/epoch - 2ms/step
Epoch 53/100
469/469 - 1s - loss: 0.5914 - sparse_categorical_accuracy: 0.7830 - val_loss: 0.4997 - val_sparse_ca
tegorical_accuracy: 0.8296 - 1s/epoch - 2ms/step
Epoch 54/100
469/469 - 1s - loss: 0.5815 - sparse_categorical_accuracy: 0.7854 - val_loss: 0.5092 - val_sparse_ca
tegorical_accuracy: 0.8220 - 1s/epoch - 2ms/step
Epoch 55/100
469/469 - 1s - loss: 0.5894 - sparse_categorical_accuracy: 0.7855 - val_loss: 0.5671 - val_sparse_ca
tegorical_accuracy: 0.8056 - 1s/epoch - 2ms/step
Epoch 56/100
469/469 - 1s - loss: 0.5820 - sparse_categorical_accuracy: 0.7853 - val_loss: 0.5121 - val_sparse_ca
tegorical_accuracy: 0.8283 - 1s/epoch - 3ms/step
Epoch 57/100
469/469 - 1s - loss: 0.5818 - sparse_categorical_accuracy: 0.7858 - val_loss: 0.5411 - val_sparse_ca
tegorical_accuracy: 0.8158 - 1s/epoch - 2ms/step
Epoch 58/100
469/469 - 1s - loss: 0.5819 - sparse_categorical_accuracy: 0.7865 - val_loss: 0.5350 - val_sparse_ca
tegorical_accuracy: 0.8222 - 1s/epoch - 2ms/step
Epoch 59/100
469/469 - 1s - loss: 0.5826 - sparse_categorical_accuracy: 0.7864 - val_loss: 0.5424 - val_sparse_ca
tegorical_accuracy: 0.8164 - 1s/epoch - 2ms/step
Epoch 60/100
469/469 - 1s - loss: 0.5859 - sparse_categorical_accuracy: 0.7855 - val_loss: 0.5348 - val_sparse_ca
tegorical_accuracy: 0.8201 - 1s/epoch - 2ms/step
Epoch 61/100
469/469 - 1s - loss: 0.5856 - sparse_categorical_accuracy: 0.7855 - val_loss: 0.5923 - val_sparse_ca
tegorical_accuracy: 0.7961 - 1s/epoch - 2ms/step
Epoch 62/100
469/469 - 1s - loss: 0.5792 - sparse_categorical_accuracy: 0.7872 - val_loss: 0.5251 - val_sparse_ca
tegorical_accuracy: 0.8223 - 1s/epoch - 2ms/step
Epoch 63/100
469/469 - 1s - loss: 0.5790 - sparse_categorical_accuracy: 0.7871 - val_loss: 0.5476 - val_sparse_ca
tegorical_accuracy: 0.8125 - 1s/epoch - 2ms/step
Epoch 64/100
469/469 - 1s - loss: 0.5797 - sparse_categorical_accuracy: 0.7879 - val_loss: 0.5142 - val_sparse_ca
tegorical_accuracy: 0.8227 - 1s/epoch - 2ms/step
Epoch 65/100
469/469 - 1s - loss: 0.5860 - sparse_categorical_accuracy: 0.7850 - val_loss: 0.5394 - val_sparse_ca
tegorical_accuracy: 0.8145 - 1s/epoch - 2ms/step
Epoch 66/100
469/469 - 1s - loss: 0.5778 - sparse_categorical_accuracy: 0.7875 - val_loss: 0.5714 - val_sparse_ca
tegorical_accuracy: 0.7929 - 1s/epoch - 2ms/step
Epoch 67/100
```

```
469/469 - 1s - loss: 0.5838 - sparse_categorical_accuracy: 0.7846 - val_loss: 0.5175 - val_sparse_ca
tegorical_accuracy: 0.8225 - 1s/epoch - 2ms/step
Epoch 68/100
469/469 - 1s - loss: 0.5777 - sparse_categorical_accuracy: 0.7893 - val_loss: 0.5637 - val_sparse_ca
tegorical_accuracy: 0.7986 - 1s/epoch - 3ms/step
Epoch 69/100
469/469 - 1s - loss: 0.5891 - sparse_categorical_accuracy: 0.7852 - val_loss: 0.5391 - val_sparse_ca
tegorical_accuracy: 0.8171 - 1s/epoch - 2ms/step
Epoch 70/100
469/469 - 1s - loss: 0.5828 - sparse_categorical_accuracy: 0.7859 - val_loss: 0.5416 - val_sparse_ca
tegorical_accuracy: 0.8152 - 1s/epoch - 2ms/step
Epoch 71/100
469/469 - 1s - loss: 0.5892 - sparse_categorical_accuracy: 0.7829 - val_loss: 0.5174 - val_sparse_ca
tegorical_accuracy: 0.8236 - 1s/epoch - 2ms/step
Epoch 72/100
469/469 - 1s - loss: 0.5846 - sparse_categorical_accuracy: 0.7848 - val_loss: 0.5145 - val_sparse_ca
tegorical_accuracy: 0.8193 - 1s/epoch - 3ms/step
Epoch 73/100
469/469 - 1s - loss: 0.5764 - sparse_categorical_accuracy: 0.7886 - val_loss: 0.5420 - val_sparse_ca
tegorical_accuracy: 0.8122 - 1s/epoch - 2ms/step
Epoch 74/100
469/469 - 1s - loss: 0.5847 - sparse_categorical_accuracy: 0.7852 - val_loss: 0.5207 - val_sparse_ca
tegorical_accuracy: 0.8213 - 1s/epoch - 2ms/step
Epoch 75/100
469/469 - 1s - loss: 0.5817 - sparse_categorical_accuracy: 0.7858 - val_loss: 0.5250 - val_sparse_ca
tegorical_accuracy: 0.8185 - 1s/epoch - 2ms/step
Epoch 76/100
469/469 - 1s - loss: 0.5754 - sparse_categorical_accuracy: 0.7892 - val_loss: 0.5776 - val_sparse_ca
tegorical_accuracy: 0.8001 - 1s/epoch - 2ms/step
Epoch 77/100
469/469 - 1s - loss: 0.5770 - sparse_categorical_accuracy: 0.7878 - val_loss: 0.5853 - val_sparse_ca
tegorical_accuracy: 0.8056 - 1s/epoch - 2ms/step
Epoch 78/100
469/469 - 1s - loss: 0.5800 - sparse_categorical_accuracy: 0.7886 - val_loss: 0.5253 - val_sparse_ca
tegorical_accuracy: 0.8192 - 1s/epoch - 2ms/step
Epoch 79/100
469/469 - 1s - loss: 0.5826 - sparse_categorical_accuracy: 0.7866 - val_loss: 0.5260 - val_sparse_ca
tegorical_accuracy: 0.8148 - 1s/epoch - 2ms/step
Epoch 80/100
469/469 - 1s - loss: 0.5743 - sparse_categorical_accuracy: 0.7883 - val_loss: 0.5056 - val_sparse_ca
tegorical_accuracy: 0.8261 - 1s/epoch - 2ms/step
Epoch 81/100
469/469 - 1s - loss: 0.5749 - sparse_categorical_accuracy: 0.7887 - val_loss: 0.5733 - val_sparse_ca
tegorical_accuracy: 0.8011 - 1s/epoch - 2ms/step
Epoch 82/100
469/469 - 1s - loss: 0.5725 - sparse_categorical_accuracy: 0.7905 - val_loss: 0.5234 - val_sparse_ca
tegorical_accuracy: 0.8196 - 1s/epoch - 2ms/step
Epoch 83/100
469/469 - 1s - loss: 0.5833 - sparse_categorical_accuracy: 0.7853 - val_loss: 0.5439 - val_sparse_ca
tegorical_accuracy: 0.8187 - 1s/epoch - 2ms/step
Epoch 84/100
469/469 - 1s - loss: 0.5824 - sparse_categorical_accuracy: 0.7841 - val_loss: 0.5209 - val_sparse_ca
tegorical_accuracy: 0.8152 - 1s/epoch - 2ms/step
Epoch 85/100
469/469 - 1s - loss: 0.5767 - sparse_categorical_accuracy: 0.7893 - val_loss: 0.5554 - val_sparse_ca
tegorical_accuracy: 0.8153 - 1s/epoch - 2ms/step
Epoch 86/100
469/469 - 1s - loss: 0.5836 - sparse_categorical_accuracy: 0.7858 - val_loss: 0.5138 - val_sparse_ca
tegorical_accuracy: 0.8195 - 1s/epoch - 2ms/step
Epoch 87/100
469/469 - 1s - loss: 0.5822 - sparse_categorical_accuracy: 0.7882 - val_loss: 0.5517 - val_sparse_ca
tegorical_accuracy: 0.8093 - 1s/epoch - 2ms/step
Epoch 88/100
469/469 - 1s - loss: 0.5805 - sparse_categorical_accuracy: 0.7884 - val_loss: 0.5915 - val_sparse_ca
tegorical_accuracy: 0.7875 - 1s/epoch - 3ms/step
Epoch 89/100
469/469 - 1s - loss: 0.5745 - sparse_categorical_accuracy: 0.7879 - val_loss: 0.5423 - val_sparse_ca
tegorical_accuracy: 0.8151 - 1s/epoch - 2ms/step
Epoch 90/100
469/469 - 1s - loss: 0.5762 - sparse_categorical_accuracy: 0.7887 - val_loss: 0.5201 - val_sparse_ca
tegorical_accuracy: 0.8238 - 1s/epoch - 3ms/step
Epoch 91/100
469/469 - 1s - loss: 0.5728 - sparse_categorical_accuracy: 0.7901 - val_loss: 0.5108 - val_sparse_ca
tegorical_accuracy: 0.8281 - 1s/epoch - 3ms/step
Epoch 92/100
469/469 - 1s - loss: 0.5806 - sparse_categorical_accuracy: 0.7862 - val_loss: 0.5754 - val_sparse_ca
tegorical_accuracy: 0.8008 - 1s/epoch - 2ms/step
Epoch 93/100
469/469 - 1s - loss: 0.5788 - sparse_categorical_accuracy: 0.7882 - val_loss: 0.5459 - val_sparse_ca
tegorical_accuracy: 0.8125 - 1s/epoch - 3ms/step
Epoch 94/100
469/469 - 1s - loss: 0.5796 - sparse_categorical_accuracy: 0.7886 - val_loss: 0.5142 - val_sparse_ca
tegorical_accuracy: 0.8270 - 1s/epoch - 2ms/step
```

```
Epoch 95/100
469/469 - 1s - loss: 0.5729 - sparse_categorical_accuracy: 0.7903 - val_loss: 0.5299 - val_sparse_ca
tegorical_accuracy: 0.8184 - 1s/epoch - 3ms/step
Epoch 96/100
469/469 - 1s - loss: 0.5715 - sparse_categorical_accuracy: 0.7889 - val_loss: 0.5318 - val_sparse_ca
tegorical_accuracy: 0.8181 - 1s/epoch - 2ms/step
Epoch 97/100
469/469 - 1s - loss: 0.5820 - sparse_categorical_accuracy: 0.7870 - val_loss: 0.5498 - val_sparse_ca
tegorical_accuracy: 0.8163 - 1s/epoch - 2ms/step
Epoch 98/100
469/469 - 1s - loss: 0.5736 - sparse_categorical_accuracy: 0.7904 - val_loss: 0.6625 - val_sparse_ca
tegorical_accuracy: 0.7919 - 1s/epoch - 2ms/step
Epoch 99/100
469/469 - 1s - loss: 0.5746 - sparse_categorical_accuracy: 0.7889 - val_loss: 0.5318 - val_sparse_ca
tegorical_accuracy: 0.8228 - 1s/epoch - 2ms/step
Epoch 100/100
469/469 - 1s - loss: 0.5716 - sparse_categorical_accuracy: 0.7902 - val_loss: 0.5275 - val_sparse_ca
tegorical_accuracy: 0.8094 - 1s/epoch - 2ms/step
```

In [122]:

```
model10_f.evaluate(testX_f, testy_f, verbose=2)
```

```
313/313 - 0s - loss: 0.5275 - sparse_categorical_accuracy: 0.8094 - 385ms/epoch - 1ms/step
```

Out[122]:

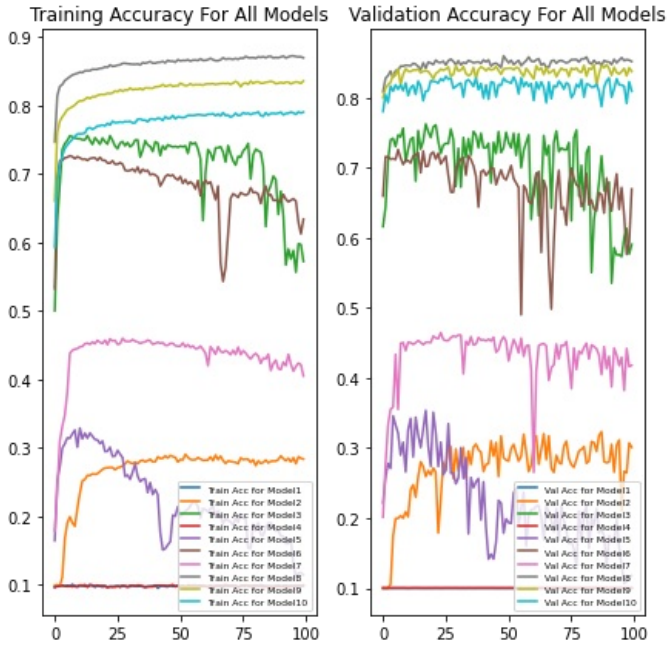
```
[0.5275380611419678, 0.8094000220298767]
```

Plots

Plots for different regularization setups

In [123]:

```
training_accuracy1_f = history1_f.history['sparse_categorical_accuracy']
validation_accuracy1_f = history1_f.history['val_sparse_categorical_accuracy']

training_accuracy2_f = history2_f.history['sparse_categorical_accuracy']
validation_accuracy2_f = history2_f.history['val_sparse_categorical_accuracy']

training_accuracy3_f = history3_f.history['sparse_categorical_accuracy']
validation_accuracy3_f = history3_f.history['val_sparse_categorical_accuracy']

training_accuracy4_f = history4_f.history['sparse_categorical_accuracy']
validation_accuracy4_f = history4_f.history['val_sparse_categorical_accuracy']

training_accuracy5_f = history5_f.history['sparse_categorical_accuracy']
validation_accuracy5_f = history5_f.history['val_sparse_categorical_accuracy']

training_accuracy6_f = history6_f.history['sparse_categorical_accuracy']
validation_accuracy6_f = history6_f.history['val_sparse_categorical_accuracy']

training_accuracy7_f = history7_f.history['sparse_categorical_accuracy']
validation_accuracy7_f = history7_f.history['val_sparse_categorical_accuracy']

training_accuracy8_f = history8_f.history['sparse_categorical_accuracy']
validation_accuracy8_f = history8_f.history['val_sparse_categorical_accuracy']

training_accuracy9_f = history9_f.history['sparse_categorical_accuracy']
validation_accuracy9_f = history9_f.history['val_sparse_categorical_accuracy']

training_accuracy10_f = history10_f.history['sparse_categorical_accuracy']
validation_accuracy10_f = history10_f.history['val_sparse_categorical_accuracy']

epochs_range=range(100)

plt.figure(figsize=(7, 7))
plt.subplot(1, 2, 1)
plt.plot(epochs_range, training_accuracy1_f, label='Train Acc for Model1')
plt.plot(epochs_range, training_accuracy2_f, label='Train Acc for Model2')
plt.plot(epochs_range, training_accuracy3_f, label='Train Acc for Model3')
plt.plot(epochs_range, training_accuracy4_f, label='Train Acc for Model4')
plt.plot(epochs_range, training_accuracy5_f, label='Train Acc for Model5')
plt.plot(epochs_range, training_accuracy6_f, label='Train Acc for Model6')
plt.plot(epochs_range, training_accuracy7_f, label='Train Acc for Model7')
plt.plot(epochs_range, training_accuracy8_f, label='Train Acc for Model8')
plt.plot(epochs_range, training_accuracy9_f, label='Train Acc for Model9')
plt.plot(epochs_range, training_accuracy10_f, label='Train Acc for Model10')
plt.legend(loc='lower right', prop={'size': 6})
plt.title('Training Accuracy For All Models')

plt.subplot(1, 2, 2)
plt.plot(epochs_range, validation_accuracy1_f, label='Val Acc for Model1')
plt.plot(epochs_range, validation_accuracy2_f, label='Val Acc for Model2')
plt.plot(epochs_range, validation_accuracy3_f, label='Val Acc for Model3')
plt.plot(epochs_range, validation_accuracy4_f, label='Val Acc for Model4')
plt.plot(epochs_range, validation_accuracy5_f, label='Val Acc for Model5')
plt.plot(epochs_range, validation_accuracy6_f, label='Val Acc for Model6')
plt.plot(epochs_range, validation_accuracy7_f, label='Val Acc for Model7')
plt.plot(epochs_range, validation_accuracy8_f, label='Val Acc for Model8')
plt.plot(epochs_range, validation_accuracy9_f, label='Val Acc for Model9')
plt.plot(epochs_range, validation_accuracy10_f, label='Val Acc for Model10')
plt.legend(loc='lower right', prop={'size': 6})
plt.title('Validation Accuracy For All Models')
plt.show()
```



Analysis of plot

Analysis:

Best model of digit dataset is model8, its out of sample accuracy can be found on the plot. And I observe the similar conclusion on fashion dataset as on digit dataset: best model is still model3. Then I can conclude that, at least in this MNIST digit/fashion datasets case, a good choice of regularization in one dataset translates into a good choice for another dataset. Which means good choices of regularization have a certain level of generalization and robustness. I believe the reason is that the idea of regularization is high level so that the performance of it wouldn't be hugely affected by detailed / individualized properties of different datasets.