白迎辰_2018201670_深度学习第三次作业

In [18]:

```
# 载入必要的函数库
import numpy as np
import matplotlib.pyplot as plt

import tensorflow as tf
from keras import Model
from keras.layers import Dense,Flatten,Input,InputLayer
from keras.optimizers import SGD,RMSprop,Adam
from keras.models import Sequential # 导入models模块中的Sequential容器

from keras.utils import to_categorical
```

In [2]:

```
## 载入mnist数据集
mnist = tf.keras.datasets.mnist
(X_train, y_train), (X_test, y_test) = mnist.load_data()

## 转换为one - hot型向量
Y_train=to_categorical(y_train)
Y_test=to_categorical(y_test)

print(Y_train.shape)
print(Y_train[0])
```

```
(60000, 10)
[0. 0. 0. 0. 0. 1. 0. 0. 0. 0.]
```

In [3]:

```
'''
 实验：构建 Multi-layer Nueral Network 模型
'''

## 第一步 创建模型结构 ##

IMSIZE = 28

input_layer = Input([IMSIZE,IMSIZE])       # MNIST图像为28*28的单层图片
x = input_layer
x = Flatten()(input_layer)                # 将28*28*1的Tensor拉直为784维向量
x = Dense(1000,activation = 'relu')(x)     # 全连接到1000个节点，并采用relu激活函数
x = Dense(10,activation = 'softmax')(x)    # 全连接到10个节点，并采用softmax激活函数转化为(0,1)取值
output_layer=x
model=Model(input_layer,output_layer)    # Model函数将input_layer 和 output_layer中间的部分连接起来

model.summary()

## 第二步 模型编译 ##

model.compile(loss='categorical_crossentropy',optimizer=SGD(lr=0.001),metrics=['accuracy'])

## 第三步 模型拟合 ##
```

```
history1 = model.fit(X_train,Y_train, validation_data=(X_test,Y_test), batch_size=1000, epochs=50)

# 第四部 提取loss指标
# model.fit会返回一个history对象，里面记录了训练集和测试集的loss以及acc
# 我们将这些指标取出，绘制折线图

train_loss1 = history1.history["loss"]
```

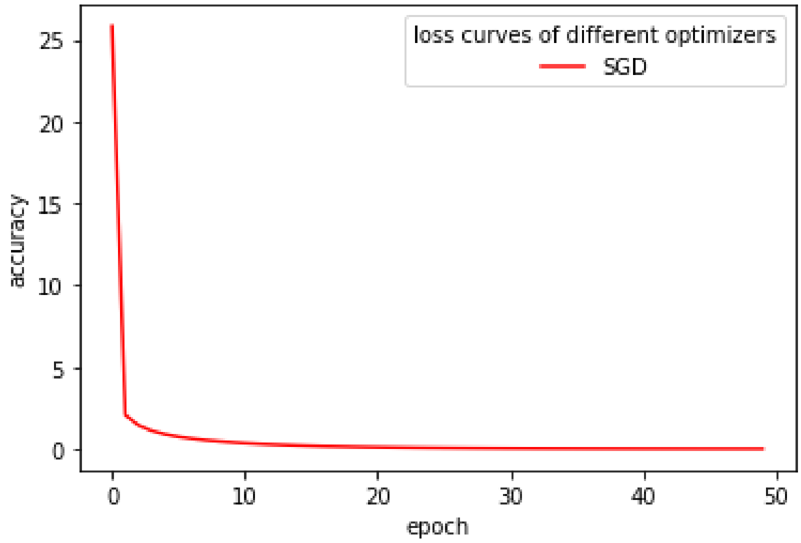
```
Model: "model_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_1 (InputLayer)         (None, 28, 28)            0
_________________________________________________________________
flatten_1 (Flatten)          (None, 784)               0
_________________________________________________________________
dense_1 (Dense)              (None, 1000)              785000
_________________________________________________________________
dense_2 (Dense)              (None, 10)                10010
=================================================================
Total params: 795,010
Trainable params: 795,010
Non-trainable params: 0
_________________________________________________________________
Train on 60000 samples, validate on 10000 samples
Epoch 1/50
60000/60000 [==============================] - 1s 8us/step - loss: 25.8237 - accuracy: 0.8042 - val_loss: 2.5986 - val_accuracy: 0.8942
Epoch 2/50
60000/60000 [==============================] - 0s 5us/step - loss: 2.0604 - accuracy: 0.9045 - val_loss: 1.8495 - val_accuracy: 0.9095
Epoch 3/50
60000/60000 [==============================] - 0s 5us/step - loss: 1.4379 - accuracy: 0.9220 - val_loss: 1.5483 - val_accuracy: 0.9181
Epoch 4/50
60000/60000 [==============================] - 0s 6us/step - loss: 1.0997 - accuracy: 0.9342 - val_loss: 1.3622 - val_accuracy: 0.9249
Epoch 5/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.8873 - accuracy: 0.9416 - val_loss: 1.2538 - val_accuracy: 0.9293
Epoch 6/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.7341 - accuracy: 0.9480 - val_loss: 1.1744 - val_accuracy: 0.9314
Epoch 7/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.6215 - accuracy: 0.9541 - val_loss: 1.1088 - val_accuracy: 0.9325
Epoch 8/50
60000/60000 [==============================] - 0s 6us/step - loss: 0.5316 - accuracy: 0.9588 - val_loss: 1.0637 - val_accuracy: 0.9351
Epoch 9/50
60000/60000 [==============================] - 0s 6us/step - loss: 0.4591 - accuracy: 0.9623 - val_loss: 1.0316 - val_accuracy: 0.9354
Epoch 10/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.3984 - accuracy: 0.9658 - val_loss: 1.0036 - val_accuracy: 0.9354
Epoch 11/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.3457 - accuracy: 0.9690 - val_loss: 0.9753 - val_accuracy: 0.9367
Epoch 12/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.3042 - accuracy: 0.9712 - val_loss: 0.9592 - val_accuracy: 0.9369
Epoch 13/50
```

```
60000/60000 [==============================] - 1s 9us/step - loss: 0.2656 - accuracy: 0.9736 - val_loss: 0.9378 - val_accuracy: 0.9357
Epoch 14/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.2343 - accuracy: 0.9758 - val_loss: 0.9219 - val_accuracy: 0.9363
Epoch 15/50
60000/60000 [==============================] - 1s 8us/step - loss: 0.2059 - accuracy: 0.9783 - val_loss: 0.9154 - val_accuracy: 0.9355
Epoch 16/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1811 - accuracy: 0.9806 - val_loss: 0.9017 - val_accuracy: 0.9376
Epoch 17/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1595 - accuracy: 0.9823 - val_loss: 0.8939 - val_accuracy: 0.9361
Epoch 18/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.1419 - accuracy: 0.9840 - val_loss: 0.8825 - val_accuracy: 0.9378
Epoch 19/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.1255 - accuracy: 0.9855 - val_loss: 0.8761 - val_accuracy: 0.9371
Epoch 20/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.1116 - accuracy: 0.9876 - val_loss: 0.8757 - val_accuracy: 0.9367
Epoch 21/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0994 - accuracy: 0.9890 - val_loss: 0.8672 - val_accuracy: 0.9381
Epoch 22/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0886 - accuracy: 0.9898 - val_loss: 0.8637 - val_accuracy: 0.9382
Epoch 23/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0789 - accuracy: 0.9909 - val_loss: 0.8577 - val_accuracy: 0.9379
Epoch 24/50
60000/60000 [==============================] - 0s 5us/step - loss: 0.0703 - accuracy: 0.9919 - val_loss: 0.8504 - val_accuracy: 0.9385
Epoch 25/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0626 - accuracy: 0.9926 - val_loss: 0.8491 - val_accuracy: 0.9380
Epoch 26/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0557 - accuracy: 0.9938 - val_loss: 0.8468 - val_accuracy: 0.9380
Epoch 27/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0495 - accuracy: 0.9944 - val_loss: 0.8497 - val_accuracy: 0.9377
Epoch 28/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0441 - accuracy: 0.9950 - val_loss: 0.8417 - val_accuracy: 0.9379
Epoch 29/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0394 - accuracy: 0.9956 - val_loss: 0.8400 - val_accuracy: 0.9378
Epoch 30/50
60000/60000 [==============================] - 0s 6us/step - loss: 0.0348 - accuracy: 0.9961 - val_loss: 0.8342 - val_accuracy: 0.9383
Epoch 31/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0311 - accuracy: 0.9966 - val_loss: 0.8338 - val_accuracy: 0.9388
Epoch 32/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0276 - accuracy: 0.9971 - val_loss: 0.8349 - val_accuracy: 0.9392
Epoch 33/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0246 - accuracy: 0.9974 - val_loss: 0.8305 - val_accuracy: 0.9392
```

```
Epoch 34/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0216 - accuracy: 0.9979 - val_loss:
0.8307 - val_accuracy: 0.9398
Epoch 35/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0194 - accuracy: 0.9979 - val_loss:
0.8273 - val_accuracy: 0.9393
Epoch 36/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0173 - accuracy: 0.9983 - val_loss:
0.8264 - val_accuracy: 0.9394
Epoch 37/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0153 - accuracy: 0.9984 - val_loss:
0.8260 - val_accuracy: 0.9387
Epoch 38/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0137 - accuracy: 0.9987 - val_loss:
0.8222 - val_accuracy: 0.9393
Epoch 39/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0125 - accuracy: 0.9988 - val_loss:
0.8223 - val_accuracy: 0.9393
Epoch 40/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0112 - accuracy: 0.9991 - val_loss:
0.8221 - val_accuracy: 0.9394
Epoch 41/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0101 - accuracy: 0.9992 - val_loss:
0.8212 - val_accuracy: 0.9393
Epoch 42/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0092 - accuracy: 0.9993 - val_loss:
0.8192 - val_accuracy: 0.9394
Epoch 43/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0083 - accuracy: 0.9994 - val_loss:
0.8177 - val_accuracy: 0.9395
Epoch 44/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0075 - accuracy: 0.9994 - val_loss:
0.8190 - val_accuracy: 0.9399
Epoch 45/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0069 - accuracy: 0.9995 - val_loss:
0.8150 - val_accuracy: 0.9399
Epoch 46/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0063 - accuracy: 0.9995 - val_loss:
0.8148 - val_accuracy: 0.9396
Epoch 47/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0058 - accuracy: 0.9996 - val_loss:
0.8154 - val_accuracy: 0.9397
Epoch 48/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0053 - accuracy: 0.9997 - val_loss:
0.8133 - val_accuracy: 0.9399
Epoch 49/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0049 - accuracy: 0.9997 - val_loss:
0.8126 - val_accuracy: 0.9398
Epoch 50/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0045 - accuracy: 0.9997 - val_loss:
0.8115 - val_accuracy: 0.9401
```

In [4]:

```
x = np.arange(50) # 生成0:49的连续整数代表epoch
y = train_loss1 # 将history对象的loss属性命名为y
plt.plot(x,y,color='red',label='SGD') # 绘制x-y散点图并用红色平滑曲线连接，图例为SGD
plt.legend(title='loss curves of different optimizers') # 设置图例标题为不同优化器下的损失曲线
plt.xlabel('epoch') # 设置横轴表示epoch
plt.ylabel('accuracy') # 设置纵轴表示accuracy
plt.show() # 绘制图像
```



In [5]:

```
optim_list = ['SGD','RMSprop','Adagrad','Adadelta','Adam','Adamax','Nadam'] # 新建一个列表来存储所有优化器的名称
color_list = ['red','orange','yellow','green','blue','purple','grey'] # 新建一个列表来储存不同优化器对应损失曲线的颜色

for i in range(7): # 生成0:6的连续整数代表7种优化器的下标
  print('Optimizer:') # 输出当前模型使用的优化器名称
  print(optim_list[i])
  IMSIZE = 28 # 设置输入图像的长宽为28像素
  model = Sequential([InputLayer([IMSIZE,IMSIZE]),Flatten(),Dense(1000,activation='relu'),Dense(10,activation='softmax')]) # 使用sequential容器配置训练模型
  model.summary() #  输出模型信息

  model.compile(loss='categorical_crossentropy',optimizer=optim_list[i],metrics=['accuracy']) # 模型编译，损失函数为交叉熵，优化器为优化器列表中的下标为i的优化器，性能度量指标为accuracy准确率

  trained_model = model.fit(X_train,Y_train, validation_data=(X_test,Y_test), batch_size=1000, epochs=50) # 训练模型，每批1000张图，一共50批
  x = np.arange(50) # 生成0:49的连续整数代表epoch
  y = trained_model.history["loss"] # 将history对象的loss属性命名为y
  plt.plot(x,y,color=color_list[i],label=optim_list[i])# 绘制当前优化器下的损失曲线，并设置对应的图例
  plt.legend(title='loss curves of different optimizers') # 设定图片标题
plt.ylim((0,1))
plt.xlabel('epoch') #  设置横坐标为epoch
plt.ylabel('accuracy') # 设置纵坐标为accuracy
plt.show() # 绘制图像
```

```
Optimizer:
SGD
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
flatten_2 (Flatten)          (None, 784)               0
_________________________________________________________________
dense_3 (Dense)              (None, 1000)              785000
_________________________________________________________________
dense_4 (Dense)              (None, 10)                10010
=================================================================
Total params: 795,010
Trainable params: 795,010
Non-trainable params: 0
_________________________________________________________________
Train on 60000 samples, validate on 10000 samples
```

```
Train on 60000 samples, validate on 10000 samples
Epoch 1/50
60000/60000 [==============================] - 0s 6us/step - loss: 397.7557 - accuracy: 0.5533 - val_lo
ss: 0.6862 - val_accuracy: 0.8235
Epoch 2/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.5725 - accuracy: 0.8522 - val_loss:
0.4978 - val_accuracy: 0.8761
Epoch 3/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.4173 - accuracy: 0.8869 - val_loss:
0.3678 - val_accuracy: 0.8987
Epoch 4/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.3163 - accuracy: 0.9125 - val_loss:
0.2968 - val_accuracy: 0.9171
Epoch 5/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.2752 - accuracy: 0.9225 - val_loss:
0.3213 - val_accuracy: 0.9117
Epoch 6/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.2407 - accuracy: 0.9316 - val_loss:
0.2921 - val_accuracy: 0.9201
Epoch 7/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.2173 - accuracy: 0.9374 - val_loss:
0.2454 - val_accuracy: 0.9320
Epoch 8/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.2092 - accuracy: 0.9395 - val_loss:
0.2326 - val_accuracy: 0.9354
Epoch 9/50
60000/60000 [==============================] - 0s 6us/step - loss: 0.1864 - accuracy: 0.9457 - val_loss:
0.2257 - val_accuracy: 0.9387
Epoch 10/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.1782 - accuracy: 0.9474 - val_loss:
0.2206 - val_accuracy: 0.9397
Epoch 11/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.1640 - accuracy: 0.9514 - val_loss:
0.2132 - val_accuracy: 0.9416
Epoch 12/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.1564 - accuracy: 0.9542 - val_loss:
0.2091 - val_accuracy: 0.9436
Epoch 13/50
60000/60000 [==============================] - 1s 8us/step - loss: 0.1479 - accuracy: 0.9557 - val_loss:
0.2058 - val_accuracy: 0.9480
Epoch 14/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1385 - accuracy: 0.9580 - val_los
s: 0.2092 - val_accuracy: 0.9439
Epoch 15/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1365 - accuracy: 0.9591 - val_los
s: 0.1954 - val_accuracy: 0.9485
Epoch 16/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1284 - accuracy: 0.9611 - val_los
s: 0.2624 - val_accuracy: 0.9296
Epoch 17/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1254 - accuracy: 0.9611 - val_los
s: 0.1967 - val_accuracy: 0.9487
Epoch 18/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1194 - accuracy: 0.9633 - val_los
s: 0.1899 - val_accuracy: 0.9503
Epoch 19/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1135 - accuracy: 0.9650 - val_los
s: 0.1959 - val_accuracy: 0.9506
Epoch 20/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1118 - accuracy: 0.9657 - val_los
s: 0.2042 - val_accuracy: 0.9468
Epoch 21/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1068 - accuracy: 0.9675 - val_los
```

```
60000/60000 [==============================] - 1s 10us/step - loss: 0.1068 - accuracy: 0.9675 - val_los
s: 0.1910 - val_accuracy: 0.9512
Epoch 22/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1049 - accuracy: 0.9680 - val_los
s: 0.1865 - val_accuracy: 0.9531
Epoch 23/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1002 - accuracy: 0.9691 - val_los
s: 0.1899 - val_accuracy: 0.9524
Epoch 24/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0975 - accuracy: 0.9700 - val_los
s: 0.1878 - val_accuracy: 0.9539
Epoch 25/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0946 - accuracy: 0.9709 - val_los
s: 0.1862 - val_accuracy: 0.9553
Epoch 26/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0923 - accuracy: 0.9719 - val_los
s: 0.1959 - val_accuracy: 0.9517
Epoch 27/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0895 - accuracy: 0.9729 - val_loss:
0.1850 - val_accuracy: 0.9566
Epoch 28/50
60000/60000 [==============================] - 1s 8us/step - loss: 0.0868 - accuracy: 0.9735 - val_loss:
0.1844 - val_accuracy: 0.9555
Epoch 29/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0864 - accuracy: 0.9738 - val_los
s: 0.1823 - val_accuracy: 0.9563
Epoch 30/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0831 - accuracy: 0.9746 - val_los
s: 0.1843 - val_accuracy: 0.9548
Epoch 31/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0837 - accuracy: 0.9747 - val_los
s: 0.1911 - val_accuracy: 0.9536
Epoch 32/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0783 - accuracy: 0.9765 - val_los
s: 0.1834 - val_accuracy: 0.9552
Epoch 33/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0763 - accuracy: 0.9769 - val_los
s: 0.1842 - val_accuracy: 0.9558
Epoch 34/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0753 - accuracy: 0.9774 - val_los
s: 0.1832 - val_accuracy: 0.9547
Epoch 35/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0734 - accuracy: 0.9775 - val_los
s: 0.1832 - val_accuracy: 0.9561
Epoch 36/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0711 - accuracy: 0.9783 - val_los
s: 0.1838 - val_accuracy: 0.9555
Epoch 37/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0696 - accuracy: 0.9790 - val_los
s: 0.1848 - val_accuracy: 0.9562
Epoch 38/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0682 - accuracy: 0.9794 - val_los
s: 0.1834 - val_accuracy: 0.9566
Epoch 39/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0669 - accuracy: 0.9801 - val_los
s: 0.1827 - val_accuracy: 0.9575
Epoch 40/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0650 - accuracy: 0.9803 - val_los
s: 0.1825 - val_accuracy: 0.9565
Epoch 41/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0638 - accuracy: 0.9805 - val_los
s: 0.1864 - val_accuracy: 0.9576
Epoch 42/50
```

```
Epoch 42/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0625 - accuracy: 0.9811 - val_los
s: 0.1840 - val_accuracy: 0.9566
Epoch 43/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0621 - accuracy: 0.9814 - val_loss:
0.1861 - val_accuracy: 0.9586
Epoch 44/50
60000/60000 [==============================] - 0s 5us/step - loss: 0.0597 - accuracy: 0.9820 - val_loss:
0.1814 - val_accuracy: 0.9594
Epoch 45/50
60000/60000 [==============================] - 0s 6us/step - loss: 0.0587 - accuracy: 0.9824 - val_loss:
0.1830 - val_accuracy: 0.9572
Epoch 46/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0568 - accuracy: 0.9830 - val_loss:
0.1853 - val_accuracy: 0.9590
Epoch 47/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0559 - accuracy: 0.9836 - val_loss:
0.1876 - val_accuracy: 0.9570
Epoch 48/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0553 - accuracy: 0.9833 - val_loss:
0.1818 - val_accuracy: 0.9594
Epoch 49/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0537 - accuracy: 0.9839 - val_loss:
0.1868 - val_accuracy: 0.9601
Epoch 50/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0525 - accuracy: 0.9845 - val_loss:
0.1837 - val_accuracy: 0.9594
Optimizer:
RMSprop
Model: "sequential_2"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
flatten_3 (Flatten)          (None, 784)               0
_________________________________________________________________
dense_5 (Dense)              (None, 1000)              785000
_________________________________________________________________
dense_6 (Dense)              (None, 10)                10010
=================================================================
Total params: 795,010
Trainable params: 795,010
Non-trainable params: 0
_________________________________________________________________
Train on 60000 samples, validate on 10000 samples
Epoch 1/50
60000/60000 [==============================] - 0s 7us/step - loss: 30.2334 - accuracy: 0.8171 - val_los
s: 5.7740 - val_accuracy: 0.8655
Epoch 2/50
60000/60000 [==============================] - 0s 8us/step - loss: 2.9646 - accuracy: 0.9208 - val_loss:
0.8965 - val_accuracy: 0.9566
Epoch 3/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.8181 - accuracy: 0.9543 - val_los
s: 0.7946 - val_accuracy: 0.9536
Epoch 4/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.4145 - accuracy: 0.9673 - val_los
s: 0.5843 - val_accuracy: 0.9603
Epoch 5/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.2854 - accuracy: 0.9746 - val_los
s: 0.4999 - val_accuracy: 0.9698
Epoch 6/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.2216 - accuracy: 0.9798 - val_los
s: 0.4961 - val_accuracy: 0.9706
Epoch 7/50
```

```
Epoch 7/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.1859 - accuracy: 0.9834 - val_los
s: 0.5392 - val_accuracy: 0.9689
Epoch 8/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.1558 - accuracy: 0.9867 - val_loss:
0.5838 - val_accuracy: 0.9708
Epoch 9/50
60000/60000 [==============================] - 1s 8us/step - loss: 0.1468 - accuracy: 0.9870 - val_loss:
0.6996 - val_accuracy: 0.9682
Epoch 10/50
60000/60000 [==============================] - 1s 8us/step - loss: 0.1344 - accuracy: 0.9886 - val_loss:
1.1087 - val_accuracy: 0.9538
Epoch 11/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.1305 - accuracy: 0.9894 - val_loss:
0.7261 - val_accuracy: 0.9717
Epoch 12/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1076 - accuracy: 0.9915 - val_los
s: 0.5916 - val_accuracy: 0.9741
Epoch 13/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0996 - accuracy: 0.9915 - val_loss:
0.6339 - val_accuracy: 0.9731
Epoch 14/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0941 - accuracy: 0.9924 - val_los
s: 0.6572 - val_accuracy: 0.9753
Epoch 15/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0916 - accuracy: 0.9926 - val_los
s: 0.6110 - val_accuracy: 0.9764
Epoch 16/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1034 - accuracy: 0.9924 - val_los
s: 0.7481 - val_accuracy: 0.9751
Epoch 17/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0864 - accuracy: 0.9932 - val_loss:
0.6473 - val_accuracy: 0.9765
Epoch 18/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0625 - accuracy: 0.9946 - val_loss:
0.7424 - val_accuracy: 0.9753
Epoch 19/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0766 - accuracy: 0.9940 - val_los
s: 0.8345 - val_accuracy: 0.9747
Epoch 20/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0901 - accuracy: 0.9936 - val_los
s: 0.7422 - val_accuracy: 0.9773
Epoch 21/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0653 - accuracy: 0.9955 - val_loss:
0.8816 - val_accuracy: 0.9726
Epoch 22/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0753 - accuracy: 0.9942 - val_loss:
0.8109 - val_accuracy: 0.9760
Epoch 23/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0483 - accuracy: 0.9962 - val_loss:
0.7151 - val_accuracy: 0.9788
Epoch 24/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0711 - accuracy: 0.9951 - val_loss:
0.8573 - val_accuracy: 0.9778
Epoch 25/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0714 - accuracy: 0.9953 - val_loss:
0.7429 - val_accuracy: 0.9796
Epoch 26/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0526 - accuracy: 0.9965 - val_los
s: 0.8046 - val_accuracy: 0.9779
Epoch 27/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0597 - accuracy: 0.9959 - val_loss:
```

```
0.7927 - val_accuracy: 0.9794
Epoch 28/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0662 - accuracy: 0.9959 - val_loss:
0.8404 - val_accuracy: 0.9786
Epoch 29/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0597 - accuracy: 0.9961 - val_loss:
0.7879 - val_accuracy: 0.9803
Epoch 30/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0479 - accuracy: 0.9968 - val_loss:
0.7560 - val_accuracy: 0.9810
Epoch 31/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0556 - accuracy: 0.9965 - val_loss:
0.8784 - val_accuracy: 0.9796
Epoch 32/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0440 - accuracy: 0.9972 - val_loss:
0.8396 - val_accuracy: 0.9792
Epoch 33/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0400 - accuracy: 0.9973 - val_los
s: 0.8777 - val_accuracy: 0.9799
Epoch 34/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0659 - accuracy: 0.9964 - val_los
s: 0.9000 - val_accuracy: 0.9803
Epoch 35/50
60000/60000 [==============================] - 1s 8us/step - loss: 0.0488 - accuracy: 0.9967 - val_loss:
0.9287 - val_accuracy: 0.9789
Epoch 36/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0478 - accuracy: 0.9973 - val_loss:
0.9804 - val_accuracy: 0.9782
Epoch 37/50
60000/60000 [==============================] - 1s 8us/step - loss: 0.0484 - accuracy: 0.9971 - val_loss:
0.9415 - val_accuracy: 0.9798
Epoch 38/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0464 - accuracy: 0.9971 - val_loss:
0.9183 - val_accuracy: 0.9797
Epoch 39/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0465 - accuracy: 0.9972 - val_loss:
0.9111 - val_accuracy: 0.9820
Epoch 40/50
60000/60000 [==============================] - 1s 8us/step - loss: 0.0413 - accuracy: 0.9974 - val_loss:
1.0465 - val_accuracy: 0.9772
Epoch 41/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0487 - accuracy: 0.9974 - val_loss:
1.3392 - val_accuracy: 0.9725
Epoch 42/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0364 - accuracy: 0.9977 - val_loss:
1.1656 - val_accuracy: 0.9770
Epoch 43/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0423 - accuracy: 0.9977 - val_los
s: 1.0539 - val_accuracy: 0.9779
Epoch 44/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0410 - accuracy: 0.9977 - val_los
s: 1.1126 - val_accuracy: 0.9790
Epoch 45/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0368 - accuracy: 0.9977 - val_los
s: 0.9085 - val_accuracy: 0.9803
Epoch 46/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0358 - accuracy: 0.9979 - val_los
s: 1.2239 - val_accuracy: 0.9765
Epoch 47/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0369 - accuracy: 0.9979 - val_los
s: 1.1020 - val_accuracy: 0.9798
Epoch 48/50
```

```
60000/60000 [==============================] - 0s 7us/step - loss: 0.0451 - accuracy: 0.9976 - val_loss: 1.0508 - val_accuracy: 0.9814
Epoch 49/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0278 - accuracy: 0.9982 - val_loss: 1.0498 - val_accuracy: 0.9800
Epoch 50/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0360 - accuracy: 0.9979 - val_loss: 1.0733 - val_accuracy: 0.9785
Optimizer:
Adagrad
Model: "sequential_3"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
flatten_4 (Flatten)          (None, 784)               0
_________________________________________________________________
dense_7 (Dense)              (None, 1000)              785000
_________________________________________________________________
dense_8 (Dense)              (None, 10)                10010
=================================================================
Total params: 795,010
Trainable params: 795,010
Non-trainable params: 0
_________________________________________________________________
Train on 60000 samples, validate on 10000 samples
Epoch 1/50
60000/60000 [==============================] - 0s 7us/step - loss: 85.7621 - accuracy: 0.8219 - val_loss: 3.7914 - val_accuracy: 0.9213
Epoch 2/50
60000/60000 [==============================] - 0s 6us/step - loss: 2.4917 - accuracy: 0.9325 - val_loss: 2.2328 - val_accuracy: 0.9287
Epoch 3/50
60000/60000 [==============================] - 0s 8us/step - loss: 1.4120 - accuracy: 0.9467 - val_loss: 1.5186 - val_accuracy: 0.9429
Epoch 4/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.9800 - accuracy: 0.9541 - val_loss: 1.2905 - val_accuracy: 0.9448
Epoch 5/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.7298 - accuracy: 0.9603 - val_loss: 1.1208 - val_accuracy: 0.9475
Epoch 6/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.5506 - accuracy: 0.9657 - val_loss: 1.0737 - val_accuracy: 0.9486
Epoch 7/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.4293 - accuracy: 0.9699 - val_loss: 0.9939 - val_accuracy: 0.9508
Epoch 8/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.3608 - accuracy: 0.9731 - val_loss: 0.9768 - val_accuracy: 0.9500
Epoch 9/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.2789 - accuracy: 0.9771 - val_loss: 0.9820 - val_accuracy: 0.9491
Epoch 10/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.2384 - accuracy: 0.9791 - val_loss: 0.8391 - val_accuracy: 0.9556
Epoch 11/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1921 - accuracy: 0.9817 - val_loss: 0.8312 - val_accuracy: 0.9565
Epoch 12/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.1600 - accuracy: 0.9837 - val_loss: 0.8104 - val_accuracy: 0.9573
Epoch 13/50
```

```
60000/60000 [==============================] - 1s 10us/step - loss: 0.1309 - accuracy: 0.9862 - val_los
s: 0.8557 - val_accuracy: 0.9524
Epoch 14/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.1115 - accuracy: 0.9874 - val_loss:
0.7833 - val_accuracy: 0.9564
Epoch 15/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0922 - accuracy: 0.9894 - val_los
s: 0.7831 - val_accuracy: 0.9556
Epoch 16/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0772 - accuracy: 0.9903 - val_loss:
0.7640 - val_accuracy: 0.9573
Epoch 17/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0673 - accuracy: 0.9910 - val_loss:
0.7615 - val_accuracy: 0.9567
Epoch 18/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0549 - accuracy: 0.9926 - val_los
s: 0.7508 - val_accuracy: 0.9573
Epoch 19/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0457 - accuracy: 0.9938 - val_los
s: 0.7374 - val_accuracy: 0.9571
Epoch 20/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0397 - accuracy: 0.9941 - val_los
s: 0.7366 - val_accuracy: 0.9568
Epoch 21/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0338 - accuracy: 0.9951 - val_los
s: 0.7216 - val_accuracy: 0.9561
Epoch 22/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0267 - accuracy: 0.9965 - val_los
s: 0.7217 - val_accuracy: 0.9574
Epoch 23/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0228 - accuracy: 0.9970 - val_los
s: 0.7202 - val_accuracy: 0.9576
Epoch 24/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0187 - accuracy: 0.9974 - val_los
s: 0.7099 - val_accuracy: 0.9579
Epoch 25/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0160 - accuracy: 0.9977 - val_los
s: 0.7229 - val_accuracy: 0.9570
Epoch 26/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0135 - accuracy: 0.9981 - val_los
s: 0.7143 - val_accuracy: 0.9580
Epoch 27/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0115 - accuracy: 0.9984 - val_los
s: 0.7277 - val_accuracy: 0.9549
Epoch 28/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0095 - accuracy: 0.9987 - val_los
s: 0.7052 - val_accuracy: 0.9573
Epoch 29/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0081 - accuracy: 0.9989 - val_los
s: 0.7165 - val_accuracy: 0.9568
Epoch 30/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0066 - accuracy: 0.9992 - val_los
s: 0.7067 - val_accuracy: 0.9566
Epoch 31/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0060 - accuracy: 0.9992 - val_los
s: 0.7034 - val_accuracy: 0.9583
Epoch 32/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0050 - accuracy: 0.9994 - val_los
s: 0.7087 - val_accuracy: 0.9568
Epoch 33/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0043 - accuracy: 0.9995 - val_loss:
0.7027 - val_accuracy: 0.9577
```

```
Epoch 34/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0037 - accuracy: 0.9995 - val_loss:
0.7024 - val_accuracy: 0.9578
Epoch 35/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0032 - accuracy: 0.9997 - val_los
s: 0.7038 - val_accuracy: 0.9573
Epoch 36/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0029 - accuracy: 0.9995 - val_los
s: 0.7028 - val_accuracy: 0.9580
Epoch 37/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0024 - accuracy: 0.9997 - val_los
s: 0.7029 - val_accuracy: 0.9577
Epoch 38/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0020 - accuracy: 0.9998 - val_los
s: 0.7022 - val_accuracy: 0.9574
Epoch 39/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0018 - accuracy: 0.9998 - val_los
s: 0.7035 - val_accuracy: 0.9571
Epoch 40/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0016 - accuracy: 0.9999 - val_los
s: 0.7055 - val_accuracy: 0.9573
Epoch 41/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0014 - accuracy: 0.9999 - val_los
s: 0.7005 - val_accuracy: 0.9580
Epoch 42/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0012 - accuracy: 0.9999 - val_los
s: 0.7044 - val_accuracy: 0.9569
Epoch 43/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0011 - accuracy: 0.9999 - val_los
s: 0.7018 - val_accuracy: 0.9577
Epoch 44/50
60000/60000 [==============================] - 1s 11us/step - loss: 9.7953e-04 - accuracy: 0.9999 - val
_loss: 0.7054 - val_accuracy: 0.9573
Epoch 45/50
60000/60000 [==============================] - 1s 11us/step - loss: 9.6424e-04 - accuracy: 0.9999 - val
_loss: 0.7048 - val_accuracy: 0.9566
Epoch 46/50
60000/60000 [==============================] - 1s 11us/step - loss: 8.3550e-04 - accuracy: 0.9999 - val
_loss: 0.7004 - val_accuracy: 0.9576
Epoch 47/50
60000/60000 [==============================] - 1s 10us/step - loss: 7.4428e-04 - accuracy: 1.0000 - val
_loss: 0.7007 - val_accuracy: 0.9576
Epoch 48/50
60000/60000 [==============================] - 0s 6us/step - loss: 6.8532e-04 - accuracy: 1.0000 - val_l
oss: 0.7003 - val_accuracy: 0.9575
Epoch 49/50
60000/60000 [==============================] - 1s 8us/step - loss: 6.3169e-04 - accuracy: 1.0000 - val_l
oss: 0.7004 - val_accuracy: 0.9576
Epoch 50/50
60000/60000 [==============================] - 1s 11us/step - loss: 5.8260e-04 - accuracy: 1.0000 - val
_loss: 0.7011 - val_accuracy: 0.9578
Optimizer:
Adadelta
Model: "sequential_4"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
flatten_5 (Flatten)          (None, 784)               0
_________________________________________________________________
dense_9 (Dense)              (None, 1000)              785000
_________________________________________________________________
dense_10 (Dense)             (None, 10)                10010
```

```
dense_10 (Dense)             (None, 10)                10010
=================================================================
Total params: 795,010
Trainable params: 795,010
Non-trainable params: 0
_________________________________________________________________
Train on 60000 samples, validate on 10000 samples
Epoch 1/50
60000/60000 [==============================] - 0s 7us/step - loss: 21.2303 - accuracy: 0.8137 - val_los
s: 8.1005 - val_accuracy: 0.8608
Epoch 2/50
60000/60000 [==============================] - 0s 8us/step - loss: 3.6146 - accuracy: 0.9197 - val_loss:
1.5753 - val_accuracy: 0.9492
Epoch 3/50
60000/60000 [==============================] - 1s 10us/step - loss: 2.4202 - accuracy: 0.9405 - val_los
s: 3.1414 - val_accuracy: 0.9017
Epoch 4/50
60000/60000 [==============================] - 1s 10us/step - loss: 1.5909 - accuracy: 0.9523 - val_los
s: 0.9695 - val_accuracy: 0.9636
Epoch 5/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.7108 - accuracy: 0.9710 - val_loss:
0.9718 - val_accuracy: 0.9565
Epoch 6/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.6918 - accuracy: 0.9697 - val_los
s: 0.8210 - val_accuracy: 0.9602
Epoch 7/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.5538 - accuracy: 0.9764 - val_los
s: 0.6200 - val_accuracy: 0.9673
Epoch 8/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.1379 - accuracy: 0.9872 - val_los
s: 0.5801 - val_accuracy: 0.9711
Epoch 9/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.3520 - accuracy: 0.9815 - val_los
s: 0.5146 - val_accuracy: 0.9722
Epoch 10/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.2129 - accuracy: 0.9851 - val_los
s: 0.4688 - val_accuracy: 0.9735
Epoch 11/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0796 - accuracy: 0.9913 - val_los
s: 0.4883 - val_accuracy: 0.9725
Epoch 12/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.2794 - accuracy: 0.9843 - val_los
s: 0.5079 - val_accuracy: 0.9723
Epoch 13/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0371 - accuracy: 0.9945 - val_los
s: 0.4494 - val_accuracy: 0.9744
Epoch 14/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0534 - accuracy: 0.9933 - val_los
s: 0.5182 - val_accuracy: 0.9751
Epoch 15/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.3441 - accuracy: 0.9856 - val_los
s: 0.4195 - val_accuracy: 0.9758
Epoch 16/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0236 - accuracy: 0.9965 - val_los
s: 0.4697 - val_accuracy: 0.9736
Epoch 17/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0452 - accuracy: 0.9942 - val_los
s: 0.4955 - val_accuracy: 0.9735
Epoch 18/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0449 - accuracy: 0.9946 - val_los
s: 0.5393 - val_accuracy: 0.9728
Epoch 19/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0329 - accuracy: 0.9958 - val_los
```

```
60000/60000 [==============================] - 1s 12us/step - loss: 0.0829 - accuracy: 0.9938 - val_los
s: 0.5560 - val_accuracy: 0.9716
Epoch 20/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0509 - accuracy: 0.9941 - val_los
s: 0.5161 - val_accuracy: 0.9768
Epoch 21/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0464 - accuracy: 0.9947 - val_los
s: 0.5281 - val_accuracy: 0.9755
Epoch 22/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0225 - accuracy: 0.9969 - val_los
s: 0.5113 - val_accuracy: 0.9760
Epoch 23/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0286 - accuracy: 0.9964 - val_los
s: 0.5459 - val_accuracy: 0.9749
Epoch 24/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0226 - accuracy: 0.9970 - val_los
s: 0.4790 - val_accuracy: 0.9780
Epoch 25/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0252 - accuracy: 0.9971 - val_los
s: 0.5384 - val_accuracy: 0.9770
Epoch 26/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0227 - accuracy: 0.9972 - val_los
s: 0.5418 - val_accuracy: 0.9770
Epoch 27/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0268 - accuracy: 0.9969 - val_los
s: 0.4966 - val_accuracy: 0.9782
Epoch 28/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0169 - accuracy: 0.9977 - val_los
s: 0.5215 - val_accuracy: 0.9788
Epoch 29/50
60000/60000 [==============================] - 0s 6us/step - loss: 0.0203 - accuracy: 0.9975 - val_loss:
0.5199 - val_accuracy: 0.9779
Epoch 30/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0194 - accuracy: 0.9978 - val_loss:
0.6079 - val_accuracy: 0.9738
Epoch 31/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0082 - accuracy: 0.9987 - val_los
s: 0.5273 - val_accuracy: 0.9767
Epoch 32/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0168 - accuracy: 0.9980 - val_los
s: 0.5011 - val_accuracy: 0.9799
Epoch 33/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0192 - accuracy: 0.9974 - val_los
s: 0.5141 - val_accuracy: 0.9789
Epoch 34/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0132 - accuracy: 0.9982 - val_los
s: 0.5388 - val_accuracy: 0.9777
Epoch 35/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0143 - accuracy: 0.9981 - val_los
s: 0.6140 - val_accuracy: 0.9770
Epoch 36/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0135 - accuracy: 0.9983 - val_los
s: 0.4927 - val_accuracy: 0.9796
Epoch 37/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0082 - accuracy: 0.9988 - val_los
s: 0.5490 - val_accuracy: 0.9778
Epoch 38/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0143 - accuracy: 0.9983 - val_los
s: 0.5913 - val_accuracy: 0.9767
Epoch 39/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0169 - accuracy: 0.9978 - val_los
s: 0.4937 - val_accuracy: 0.9789
Epoch 40/50
```

```
Epoch 40/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0179 - accuracy: 0.9980 - val_loss: 0.5282 - val_accuracy: 0.9812
Epoch 41/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0116 - accuracy: 0.9985 - val_loss: 0.5557 - val_accuracy: 0.9790
Epoch 42/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0135 - accuracy: 0.9982 - val_loss: 0.5164 - val_accuracy: 0.9784
Epoch 43/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0084 - accuracy: 0.9990 - val_loss: 0.5315 - val_accuracy: 0.9798
Epoch 44/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0097 - accuracy: 0.9986 - val_loss: 0.5254 - val_accuracy: 0.9794
Epoch 45/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0079 - accuracy: 0.9990 - val_loss: 0.5646 - val_accuracy: 0.9791
Epoch 46/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0087 - accuracy: 0.9990 - val_loss: 0.5115 - val_accuracy: 0.9800
Epoch 47/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0084 - accuracy: 0.9989 - val_loss: 0.5442 - val_accuracy: 0.9798
Epoch 48/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0056 - accuracy: 0.9993 - val_loss: 0.5115 - val_accuracy: 0.9810
Epoch 49/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0068 - accuracy: 0.9990 - val_loss: 0.5313 - val_accuracy: 0.9797
Epoch 50/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0053 - accuracy: 0.9991 - val_loss: 0.5219 - val_accuracy: 0.9809
Optimizer:
Adam
Model: "sequential_5"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
flatten_6 (Flatten)          (None, 784)               0
_________________________________________________________________
dense_11 (Dense)             (None, 1000)              785000
_________________________________________________________________
dense_12 (Dense)             (None, 10)                10010
=================================================================
Total params: 795,010
Trainable params: 795,010
Non-trainable params: 0
_________________________________________________________________
Train on 60000 samples, validate on 10000 samples
Epoch 1/50
60000/60000 [==============================] - 1s 9us/step - loss: 21.9726 - accuracy: 0.8266 - val_loss: 2.9179 - val_accuracy: 0.9319
Epoch 2/50
60000/60000 [==============================] - 0s 6us/step - loss: 1.7914 - accuracy: 0.9428 - val_loss: 1.4070 - val_accuracy: 0.9466
Epoch 3/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.7879 - accuracy: 0.9601 - val_loss: 1.0741 - val_accuracy: 0.9521
Epoch 4/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.4196 - accuracy: 0.9710 - val_loss: 0.9146 - val_accuracy: 0.9539
```

```
Epoch 5/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.2478 - accuracy: 0.9785 - val_los
s: 0.8119 - val_accuracy: 0.9577
Epoch 6/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1306 - accuracy: 0.9854 - val_los
s: 0.7619 - val_accuracy: 0.9607
Epoch 7/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0805 - accuracy: 0.9887 - val_los
s: 0.7128 - val_accuracy: 0.9646
Epoch 8/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0496 - accuracy: 0.9919 - val_los
s: 0.7085 - val_accuracy: 0.9615
Epoch 9/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0344 - accuracy: 0.9939 - val_los
s: 0.7737 - val_accuracy: 0.9624
Epoch 10/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0307 - accuracy: 0.9945 - val_loss:
0.6888 - val_accuracy: 0.9649
Epoch 11/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0246 - accuracy: 0.9951 - val_los
s: 0.7269 - val_accuracy: 0.9635
Epoch 12/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0187 - accuracy: 0.9961 - val_los
s: 0.7246 - val_accuracy: 0.9650
Epoch 13/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0185 - accuracy: 0.9968 - val_los
s: 0.6914 - val_accuracy: 0.9651
Epoch 14/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0180 - accuracy: 0.9964 - val_los
s: 0.7690 - val_accuracy: 0.9643
Epoch 15/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0161 - accuracy: 0.9969 - val_los
s: 0.7368 - val_accuracy: 0.9653
Epoch 16/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0171 - accuracy: 0.9965 - val_los
s: 0.6810 - val_accuracy: 0.9664
Epoch 17/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0181 - accuracy: 0.9965 - val_los
s: 0.6963 - val_accuracy: 0.9661
Epoch 18/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0193 - accuracy: 0.9965 - val_los
s: 0.7146 - val_accuracy: 0.9675
Epoch 19/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0203 - accuracy: 0.9966 - val_los
s: 0.7408 - val_accuracy: 0.9658
Epoch 20/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0280 - accuracy: 0.9960 - val_los
s: 0.7480 - val_accuracy: 0.9655
Epoch 21/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0307 - accuracy: 0.9957 - val_los
s: 0.7109 - val_accuracy: 0.9664
Epoch 22/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0384 - accuracy: 0.9942 - val_los
s: 0.7811 - val_accuracy: 0.9666
Epoch 23/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0334 - accuracy: 0.9947 - val_loss:
0.7773 - val_accuracy: 0.9666
Epoch 24/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0345 - accuracy: 0.9958 - val_loss:
0.7664 - val_accuracy: 0.9684
Epoch 25/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0507 - accuracy: 0.9944 - val_loss:
```

```
0.7604 - val_accuracy: 0.9672
Epoch 26/50
60000/60000 [==============================] - 0s 6us/step - loss: 0.0411 - accuracy: 0.9949 - val_loss:
0.7110 - val_accuracy: 0.9707
Epoch 27/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0483 - accuracy: 0.9939 - val_loss:
0.7508 - val_accuracy: 0.9678
Epoch 28/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0378 - accuracy: 0.9953 - val_los
s: 0.8248 - val_accuracy: 0.9697
Epoch 29/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0325 - accuracy: 0.9961 - val_loss:
0.7281 - val_accuracy: 0.9692
Epoch 30/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0337 - accuracy: 0.9957 - val_los
s: 0.6960 - val_accuracy: 0.9709
Epoch 31/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0322 - accuracy: 0.9962 - val_los
s: 0.7387 - val_accuracy: 0.9721
Epoch 32/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0312 - accuracy: 0.9958 - val_loss:
0.7564 - val_accuracy: 0.9727
Epoch 33/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0326 - accuracy: 0.9961 - val_loss:
0.7740 - val_accuracy: 0.9713
Epoch 34/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0275 - accuracy: 0.9966 - val_los
s: 0.7451 - val_accuracy: 0.9713
Epoch 35/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0338 - accuracy: 0.9959 - val_los
s: 0.7864 - val_accuracy: 0.9719
Epoch 36/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0432 - accuracy: 0.9954 - val_los
s: 0.7195 - val_accuracy: 0.9700
Epoch 37/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0253 - accuracy: 0.9970 - val_los
s: 0.7436 - val_accuracy: 0.9726
Epoch 38/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0185 - accuracy: 0.9975 - val_los
s: 0.6843 - val_accuracy: 0.9737
Epoch 39/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0225 - accuracy: 0.9970 - val_los
s: 0.7222 - val_accuracy: 0.9720
Epoch 40/50
60000/60000 [==============================] - 1s 12us/step - loss: 0.0292 - accuracy: 0.9965 - val_los
s: 0.7086 - val_accuracy: 0.9753
Epoch 41/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0290 - accuracy: 0.9965 - val_loss:
0.7089 - val_accuracy: 0.9735
Epoch 42/50
60000/60000 [==============================] - 0s 5us/step - loss: 0.0271 - accuracy: 0.9971 - val_loss:
0.7707 - val_accuracy: 0.9712
Epoch 43/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0294 - accuracy: 0.9965 - val_los
s: 0.7301 - val_accuracy: 0.9739
Epoch 44/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0326 - accuracy: 0.9965 - val_los
s: 0.7107 - val_accuracy: 0.9750
Epoch 45/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0337 - accuracy: 0.9965 - val_los
s: 0.7106 - val_accuracy: 0.9741
Epoch 46/50
```

```
60000/60000 [==============================] - 1s 11us/step - loss: 0.0392 - accuracy: 0.9959 - val_los
s: 0.8351 - val_accuracy: 0.9705
Epoch 47/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0497 - accuracy: 0.9951 - val_los
s: 0.7328 - val_accuracy: 0.9740
Epoch 48/50
60000/60000 [==============================] - 1s 11us/step - loss: 0.0388 - accuracy: 0.9965 - val_los
s: 0.7592 - val_accuracy: 0.9745
Epoch 49/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0410 - accuracy: 0.9963 - val_loss:
0.6937 - val_accuracy: 0.9743
Epoch 50/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0311 - accuracy: 0.9968 - val_loss:
0.8517 - val_accuracy: 0.9694
Optimizer:
Adamax
Model: "sequential_6"
_________________________________________________________________
Layer (type)                 Output Shape              Param #   
=================================================================
flatten_7 (Flatten)          (None, 784)               0         
_________________________________________________________________
dense_13 (Dense)             (None, 1000)              785000    
_________________________________________________________________
dense_14 (Dense)             (None, 10)                10010     
=================================================================
Total params: 795,010
Trainable params: 795,010
Non-trainable params: 0
_________________________________________________________________
Train on 60000 samples, validate on 10000 samples
Epoch 1/50
60000/60000 [==============================] - 0s 6us/step - loss: 23.0998 - accuracy: 0.8312 - val_los
s: 4.8528 - val_accuracy: 0.9223
Epoch 2/50
60000/60000 [==============================] - 0s 6us/step - loss: 3.1628 - accuracy: 0.9324 - val_loss:
2.6672 - val_accuracy: 0.9376
Epoch 3/50
60000/60000 [==============================] - 0s 8us/step - loss: 1.7192 - accuracy: 0.9479 - val_loss:
1.9301 - val_accuracy: 0.9442
Epoch 4/50
60000/60000 [==============================] - 1s 8us/step - loss: 1.0826 - accuracy: 0.9572 - val_loss:
1.6077 - val_accuracy: 0.9473
Epoch 5/50
60000/60000 [==============================] - 1s 8us/step - loss: 0.7608 - accuracy: 0.9636 - val_loss:
1.3715 - val_accuracy: 0.9499
Epoch 6/50
60000/60000 [==============================] - 1s 8us/step - loss: 0.5414 - accuracy: 0.9691 - val_loss:
1.1693 - val_accuracy: 0.9533
Epoch 7/50
60000/60000 [==============================] - 1s 8us/step - loss: 0.4033 - accuracy: 0.9733 - val_loss:
1.0628 - val_accuracy: 0.9556
Epoch 8/50
60000/60000 [==============================] - 0s 5us/step - loss: 0.3043 - accuracy: 0.9781 - val_loss:
0.9966 - val_accuracy: 0.9557
Epoch 9/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.2326 - accuracy: 0.9810 - val_loss:
0.9581 - val_accuracy: 0.9560
Epoch 10/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.1695 - accuracy: 0.9846 - val_los
s: 0.9175 - val_accuracy: 0.9566
Epoch 11/50
```

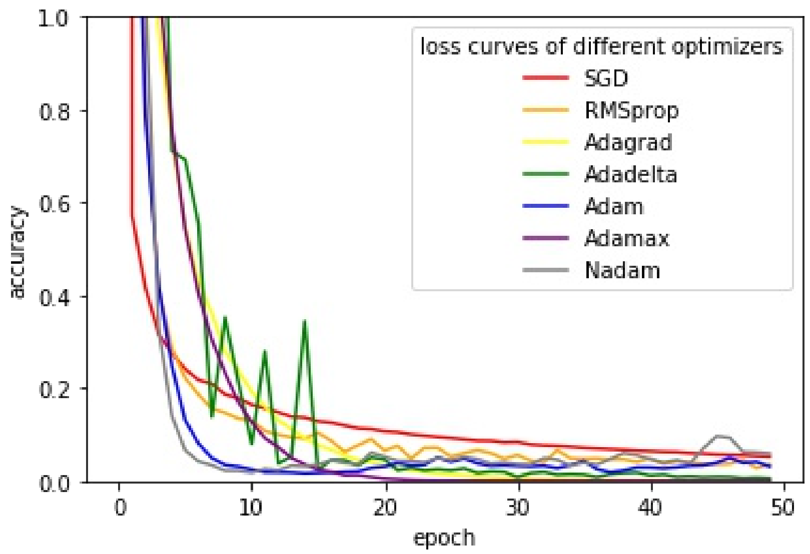
```
60000/60000 [==============================] - 0s 6us/step - loss: 0.1267 - accuracy: 0.9876 - val_loss:
0.8782 - val_accuracy: 0.9568
Epoch 12/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0923 - accuracy: 0.9897 - val_loss:
0.8753 - val_accuracy: 0.9559
Epoch 13/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0726 - accuracy: 0.9906 - val_los
s: 0.8689 - val_accuracy: 0.9560
Epoch 14/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0495 - accuracy: 0.9931 - val_los
s: 0.8472 - val_accuracy: 0.9574
Epoch 15/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0356 - accuracy: 0.9946 - val_los
s: 0.8385 - val_accuracy: 0.9583
Epoch 16/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0242 - accuracy: 0.9962 - val_los
s: 0.8087 - val_accuracy: 0.9582
Epoch 17/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0197 - accuracy: 0.9962 - val_los
s: 0.8172 - val_accuracy: 0.9578
Epoch 18/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0119 - accuracy: 0.9979 - val_los
s: 0.8084 - val_accuracy: 0.9586
Epoch 19/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0122 - accuracy: 0.9976 - val_los
s: 0.8057 - val_accuracy: 0.9598
Epoch 20/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0082 - accuracy: 0.9982 - val_los
s: 0.8002 - val_accuracy: 0.9600
Epoch 21/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0047 - accuracy: 0.9989 - val_los
s: 0.7857 - val_accuracy: 0.9601
Epoch 22/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0033 - accuracy: 0.9992 - val_los
s: 0.8062 - val_accuracy: 0.9586
Epoch 23/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0017 - accuracy: 0.9996 - val_los
s: 0.7907 - val_accuracy: 0.9596
Epoch 24/50
60000/60000 [==============================] - 1s 10us/step - loss: 8.1394e-04 - accuracy: 0.9998 - val
_loss: 0.7882 - val_accuracy: 0.9608
Epoch 25/50
60000/60000 [==============================] - 1s 10us/step - loss: 4.0861e-04 - accuracy: 1.0000 - val
_loss: 0.7860 - val_accuracy: 0.9601
Epoch 26/50
60000/60000 [==============================] - 1s 10us/step - loss: 2.1505e-04 - accuracy: 1.0000 - val
_loss: 0.7876 - val_accuracy: 0.9600
Epoch 27/50
60000/60000 [==============================] - 1s 10us/step - loss: 1.4673e-04 - accuracy: 1.0000 - val
_loss: 0.7860 - val_accuracy: 0.9604
Epoch 28/50
60000/60000 [==============================] - 1s 9us/step - loss: 1.1398e-04 - accuracy: 1.0000 - val_l
oss: 0.7898 - val_accuracy: 0.9601
Epoch 29/50
60000/60000 [==============================] - 0s 7us/step - loss: 7.1895e-05 - accuracy: 1.0000 - val_l
oss: 0.7856 - val_accuracy: 0.9605
Epoch 30/50
60000/60000 [==============================] - 0s 7us/step - loss: 5.0138e-05 - accuracy: 1.0000 - val_l
oss: 0.7869 - val_accuracy: 0.9606
Epoch 31/50
60000/60000 [==============================] - 0s 7us/step - loss: 4.2168e-05 - accuracy: 1.0000 - val_l
oss: 0.7860 - val_accuracy: 0.9604
```

```
Epoch 32/50
60000/60000 [==============================] - 0s 8us/step - loss: 3.9286e-05 - accuracy: 1.0000 - val_l
oss: 0.7864 - val_accuracy: 0.9607
Epoch 33/50
60000/60000 [==============================] - 1s 9us/step - loss: 3.6828e-05 - accuracy: 1.0000 - val_l
oss: 0.7859 - val_accuracy: 0.9605
Epoch 34/50
60000/60000 [==============================] - 1s 9us/step - loss: 3.4640e-05 - accuracy: 1.0000 - val_l
oss: 0.7856 - val_accuracy: 0.9607
Epoch 35/50
60000/60000 [==============================] - 1s 9us/step - loss: 3.2821e-05 - accuracy: 1.0000 - val_l
oss: 0.7857 - val_accuracy: 0.9605
Epoch 36/50
60000/60000 [==============================] - 1s 9us/step - loss: 3.1370e-05 - accuracy: 1.0000 - val_l
oss: 0.7857 - val_accuracy: 0.9604
Epoch 37/50
60000/60000 [==============================] - 1s 9us/step - loss: 2.9748e-05 - accuracy: 1.0000 - val_l
oss: 0.7859 - val_accuracy: 0.9604
Epoch 38/50
60000/60000 [==============================] - 1s 9us/step - loss: 2.8286e-05 - accuracy: 1.0000 - val_l
oss: 0.7852 - val_accuracy: 0.9604
Epoch 39/50
60000/60000 [==============================] - 1s 9us/step - loss: 2.7034e-05 - accuracy: 1.0000 - val_l
oss: 0.7854 - val_accuracy: 0.9604
Epoch 40/50
60000/60000 [==============================] - 1s 9us/step - loss: 2.5759e-05 - accuracy: 1.0000 - val_l
oss: 0.7852 - val_accuracy: 0.9606
Epoch 41/50
60000/60000 [==============================] - 1s 9us/step - loss: 2.4660e-05 - accuracy: 1.0000 - val_l
oss: 0.7852 - val_accuracy: 0.9605
Epoch 42/50
60000/60000 [==============================] - 1s 9us/step - loss: 2.3462e-05 - accuracy: 1.0000 - val_l
oss: 0.7848 - val_accuracy: 0.9605
Epoch 43/50
60000/60000 [==============================] - 1s 8us/step - loss: 2.2491e-05 - accuracy: 1.0000 - val_l
oss: 0.7847 - val_accuracy: 0.9606
Epoch 44/50
60000/60000 [==============================] - 1s 10us/step - loss: 2.1503e-05 - accuracy: 1.0000 - val
_loss: 0.7849 - val_accuracy: 0.9606
Epoch 45/50
60000/60000 [==============================] - 1s 10us/step - loss: 2.0657e-05 - accuracy: 1.0000 - val
_loss: 0.7847 - val_accuracy: 0.9606
Epoch 46/50
60000/60000 [==============================] - 1s 9us/step - loss: 1.9829e-05 - accuracy: 1.0000 - val_l
oss: 0.7848 - val_accuracy: 0.9607
Epoch 47/50
60000/60000 [==============================] - 1s 9us/step - loss: 1.8919e-05 - accuracy: 1.0000 - val_l
oss: 0.7848 - val_accuracy: 0.9606
Epoch 48/50
60000/60000 [==============================] - 1s 9us/step - loss: 1.8215e-05 - accuracy: 1.0000 - val_l
oss: 0.7841 - val_accuracy: 0.9608
Epoch 49/50
60000/60000 [==============================] - 1s 9us/step - loss: 1.7553e-05 - accuracy: 1.0000 - val_l
oss: 0.7843 - val_accuracy: 0.9606
Epoch 50/50
60000/60000 [==============================] - 1s 9us/step - loss: 1.6829e-05 - accuracy: 1.0000 - val_l
oss: 0.7842 - val_accuracy: 0.9605
Optimizer:
Nadam
Model: "sequential_7"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
```

```
Layer (type)                 Output Shape              Param #
=================================================================
flatten_8 (Flatten)          (None, 784)               0
_________________________________________________________________
dense_15 (Dense)             (None, 1000)              785000
_________________________________________________________________
dense_16 (Dense)             (None, 10)                10010
=================================================================
Total params: 795,010
Trainable params: 795,010
Non-trainable params: 0
_________________________________________________________________
Train on 60000 samples, validate on 10000 samples
Epoch 1/50
60000/60000 [==============================] - 0s 7us/step - loss: 19.0211 - accuracy: 0.8488 - val_los
s: 2.1557 - val_accuracy: 0.9373
Epoch 2/50
60000/60000 [==============================] - 0s 6us/step - loss: 1.7517 - accuracy: 0.9445 - val_loss:
1.4074 - val_accuracy: 0.9462
Epoch 3/50
60000/60000 [==============================] - 1s 10us/step - loss: 1.0648 - accuracy: 0.9623 - val_los
s: 1.1759 - val_accuracy: 0.9531
Epoch 4/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.3283 - accuracy: 0.9775 - val_los
s: 0.6907 - val_accuracy: 0.9628
Epoch 5/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.1401 - accuracy: 0.9864 - val_loss:
0.5747 - val_accuracy: 0.9693
Epoch 6/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0659 - accuracy: 0.9916 - val_loss:
0.6096 - val_accuracy: 0.9686
Epoch 7/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0421 - accuracy: 0.9936 - val_loss:
0.6306 - val_accuracy: 0.9671
Epoch 8/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0338 - accuracy: 0.9950 - val_loss:
0.6069 - val_accuracy: 0.9683
Epoch 9/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0219 - accuracy: 0.9966 - val_loss:
0.6243 - val_accuracy: 0.9691
Epoch 10/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0218 - accuracy: 0.9965 - val_loss:
0.6369 - val_accuracy: 0.9711
Epoch 11/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0178 - accuracy: 0.9967 - val_loss:
0.6736 - val_accuracy: 0.9664
Epoch 12/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0260 - accuracy: 0.9962 - val_loss:
0.5899 - val_accuracy: 0.9713
Epoch 13/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0230 - accuracy: 0.9962 - val_loss:
0.5925 - val_accuracy: 0.9708
Epoch 14/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0329 - accuracy: 0.9955 - val_loss:
0.6231 - val_accuracy: 0.9687
Epoch 15/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0315 - accuracy: 0.9952 - val_loss:
0.6032 - val_accuracy: 0.9709
Epoch 16/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0352 - accuracy: 0.9952 - val_loss:
0.6412 - val_accuracy: 0.9705
Epoch 17/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0445 - accuracy: 0.9949 - val_loss:
```

0.6221 - val_accuracy: 0.9724
Epoch 18/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0387 - accuracy: 0.9949 - val_loss: 0.6241 - val_accuracy: 0.9707
Epoch 19/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0355 - accuracy: 0.9956 - val_loss: 0.6430 - val_accuracy: 0.9709
Epoch 20/50
60000/60000 [==============================] - 0s 6us/step - loss: 0.0608 - accuracy: 0.9935 - val_loss: 0.7044 - val_accuracy: 0.9716
Epoch 21/50
60000/60000 [==============================] - 0s 6us/step - loss: 0.0525 - accuracy: 0.9938 - val_loss: 0.6400 - val_accuracy: 0.9708
Epoch 22/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0431 - accuracy: 0.9952 - val_loss: 0.6576 - val_accuracy: 0.9729
Epoch 23/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0429 - accuracy: 0.9952 - val_loss: 0.5967 - val_accuracy: 0.9742
Epoch 24/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0401 - accuracy: 0.9953 - val_loss: 0.6807 - val_accuracy: 0.9733
Epoch 25/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0493 - accuracy: 0.9948 - val_loss: 0.7000 - val_accuracy: 0.9735
Epoch 26/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0450 - accuracy: 0.9952 - val_loss: 0.6769 - val_accuracy: 0.9745
Epoch 27/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0518 - accuracy: 0.9951 - val_loss: 0.6583 - val_accuracy: 0.9744
Epoch 28/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0469 - accuracy: 0.9952 - val_loss: 0.6210 - val_accuracy: 0.9744
Epoch 29/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0374 - accuracy: 0.9965 - val_loss: 0.7284 - val_accuracy: 0.9728
Epoch 30/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0399 - accuracy: 0.9957 - val_loss: 0.6707 - val_accuracy: 0.9725
Epoch 31/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0361 - accuracy: 0.9963 - val_loss: 0.6618 - val_accuracy: 0.9751
Epoch 32/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0356 - accuracy: 0.9960 - val_loss: 0.6382 - val_accuracy: 0.9766
Epoch 33/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0492 - accuracy: 0.9959 - val_loss: 0.9157 - val_accuracy: 0.9676
Epoch 34/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0458 - accuracy: 0.9956 - val_loss: 0.6474 - val_accuracy: 0.9779
Epoch 35/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0327 - accuracy: 0.9968 - val_loss: 0.6614 - val_accuracy: 0.9767
Epoch 36/50
60000/60000 [==============================] - 1s 10us/step - loss: 0.0396 - accuracy: 0.9961 - val_loss: 0.6453 - val_accuracy: 0.9753
Epoch 37/50
60000/60000 [==============================] - 0s 6us/step - loss: 0.0352 - accuracy: 0.9964 - val_loss: 0.7098 - val_accuracy: 0.9754
Epoch 38/50

```
Epoch 38/50
60000/60000 [==============================] - 0s 6us/step - loss: 0.0474 - accuracy: 0.9958 - val_loss: 0.7396 - val_accuracy: 0.9727
Epoch 39/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0596 - accuracy: 0.9949 - val_loss: 0.7679 - val_accuracy: 0.9742
Epoch 40/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0556 - accuracy: 0.9957 - val_loss: 0.6952 - val_accuracy: 0.9759
Epoch 41/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0468 - accuracy: 0.9964 - val_loss: 0.7414 - val_accuracy: 0.9728
Epoch 42/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0396 - accuracy: 0.9962 - val_loss: 0.6957 - val_accuracy: 0.9761
Epoch 43/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0450 - accuracy: 0.9961 - val_loss: 0.7169 - val_accuracy: 0.9761
Epoch 44/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0376 - accuracy: 0.9968 - val_loss: 0.7498 - val_accuracy: 0.9759
Epoch 45/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0698 - accuracy: 0.9951 - val_loss: 0.9646 - val_accuracy: 0.9728
Epoch 46/50
60000/60000 [==============================] - 0s 7us/step - loss: 0.0964 - accuracy: 0.9937 - val_loss: 1.0710 - val_accuracy: 0.9653
Epoch 47/50
60000/60000 [==============================] - 0s 8us/step - loss: 0.0930 - accuracy: 0.9939 - val_loss: 0.7609 - val_accuracy: 0.9760
Epoch 48/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0622 - accuracy: 0.9953 - val_loss: 0.8329 - val_accuracy: 0.9740
Epoch 49/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0615 - accuracy: 0.9956 - val_loss: 0.7712 - val_accuracy: 0.9778
Epoch 50/50
60000/60000 [==============================] - 1s 9us/step - loss: 0.0588 - accuracy: 0.9956 - val_loss: 0.8836 - val_accuracy: 0.9751
```



深度学习优化算法大概经历了 SGD -> SGDM -> NAG ->Adagrad -> Adadelta(RMSprop) -> Adam -> Nadam 这样的发展历程。

在keras.optimizers子模块中，它们基本上都有对应的类的实现。

In [17]:

```
import keras
```

- SGD, 默认参数为纯SGD, 设置momentum参数不为0实际上变成SGDM, 考虑了一阶动量, 设置 nesterov为True后变成NAG，即 Nesterov Accelerated Gradient，在计算梯度时计算的是向前走一步所在位置的梯度。

In [10]:

```
keras.optimizers.SGD(lr=0.01, momentum=0.0, decay=0.0, nesterov=False)
```

Out[10]:

<keras.optimizers.SGD at 0x7ff3b56bebd0>

- Adagrad, 考虑了二阶动量，对于不同的参数有不同的学习率，即自适应学习率。缺点是学习率单调下降，可能后期学习速率过慢乃至提前停止学习。

In [11]:

```
keras.optimizers.Adagrad(lr=0.01, epsilon=None, decay=0.0)
```

Out[11]:

<keras.optimizers.Adagrad at 0x7ff3b56bec10>

- RMSprop, 考虑了二阶动量，对于不同的参数有不同的学习率，即自适应学习率，对Adagrad进行了优化，通过指数平滑只考虑一定窗口内的二阶动量。

In [12]:

```
keras.optimizers.RMSprop(lr=0.001, rho=0.9, epsilon=None, decay=0.0)
```

Out[12]:

<keras.optimizers.RMSprop at 0x7ff3b558d290>

- Adadelta, 考虑了二阶动量，与RMSprop类似，但是更加复杂一些，自适应性更强。

In [13]:

```
keras.optimizers.Adadelta(lr=1.0, rho=0.95, epsilon=None, decay=0.0)
```

Out[13]:

<keras.optimizers.Adadelta at 0x7ff3b558db90>

- Adam, 同时考虑了一阶动量和二阶动量，可以看成RMSprop上进一步考虑了一阶动量。

In [14]:

```
keras.optimizers.Adam(lr=0.001, beta_1=0.9, beta_2=0.999, epsilon=None, decay=0.0, amsgrad=False)
```

Out[14]:

<keras.optimizers.Adam at 0x7ff3b558e350>

- Adamax, 它是Adam算法基于无穷范数（infinity norm）的变种。

In [15]:

```
keras.optimizers.Adamax(lr=0.002, beta_1=0.9, beta_2=0.999, epsilon=None, decay=0.0)
```

Out[15]:

<keras.optimizers.Adamax at 0x7ff3b558eb90>

- Nadam, 在Adam基础上进一步考虑了 Nesterov Acceleration。

In [16]:

```
keras.optimizers.Nadam(lr=0.002, beta_1=0.9, beta_2=0.999, epsilon=None, schedule_decay=0.004)
```

Out[16]:

<keras.optimizers.Nadam at 0x7ff3b559cad0>