

姓名：张英凯 学号：3120220996

```
In [ ]: import gzip
import numpy as np
from sklearn.svm import SVC
from sklearn.metrics import accuracy_score
```

```
In [ ]: # 读取图像数据
def load_images(filename):
    with gzip.open(filename, 'rb') as f:
        num_images = int.from_bytes(f.read(4), 'big')
        num_rows = int.from_bytes(f.read(4), 'big')
        num_cols = int.from_bytes(f.read(4), 'big')
        images = np.frombuffer(f.read(), dtype=np.uint8)
        images = images.reshape(num_images, num_rows * num_cols)
    return images/255.0
```

```
In [ ]: # 读取标签数据
def load_labels(filename):
    with gzip.open(filename, 'rb') as f:
        labels = np.frombuffer(f.read(), dtype=np.uint8)
    return labels
```

```
In [ ]: # 加载训练集数据和标签
X_train = load_images('./train-images-idx3-ubyte.gz')
y_train = load_labels('./train-labels-idx1-ubyte.gz')
```

```
In [ ]: # 加载测试集数据和标签
X_test = load_images('./t10k-images-idx3-ubyte.gz')
y_test = load_labels('./t10k-labels-idx1-ubyte.gz')
```

```
In [ ]: # 打印数据集的形状
print('X_train shape:', X_train.shape)
print('y_train shape:', y_train.shape)
print('X_test shape:', X_test.shape)
print('y_test shape:', y_test.shape)
svm = SVC(kernel='linear', C=1.0, random_state=1)
svm.fit(X_train, y_train)
```

```
X_train shape: (60000, 784)
y_train shape: (60000,)
X_test shape: (10000, 784)
y_test shape: (10000,)
```

```
Out [ ]: ▼ SVC
SVC(kernel='linear', random_state=1)
```

```
In [ ]: # 在测试集上进行预测并计算准确率
y_pred = svm.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
print('Accuracy:', accuracy)
```

```
Accuracy: 0.9404
```