

# Đề kiểm tra lập trình nhập môn phân tích dữ liệu và học sâu

Sinh viên không được phép sử dụng internet

Sinh viên sau khi làm bài xong xuất ra file PDF đồng thời nộp lên Fit-lab và push lên git-hub

Sinh viên làm bắt đầu làm bài từ 15h40 - 18h00

```
In [147... # Họ và Tên: Nguyễn Minh Huy  
# MSSV: 2174802010934
```

```
In [147... import torch  
import torch.nn as nn  
import torch.optim as optim  
from torch.utils.data import Dataset, DataLoader, random_split  
import numpy as np  
from sklearn.datasets import load_iris  
import matplotlib.pyplot as plt
```

```
In [148... # Bước 1: Load data  
def load_dataset():  
    X, y = load_iris(return_X_y=True)  
    # Loại bỏ các mẫu thuộc lớp 2  
    X = X[y != 2]  
    y = y[y != 2]  
    return X, y
```

```
# Điền ở đây  
X, y = load_dataset()  
print("X:", X.shape)  
print("y:", y.shape)
```

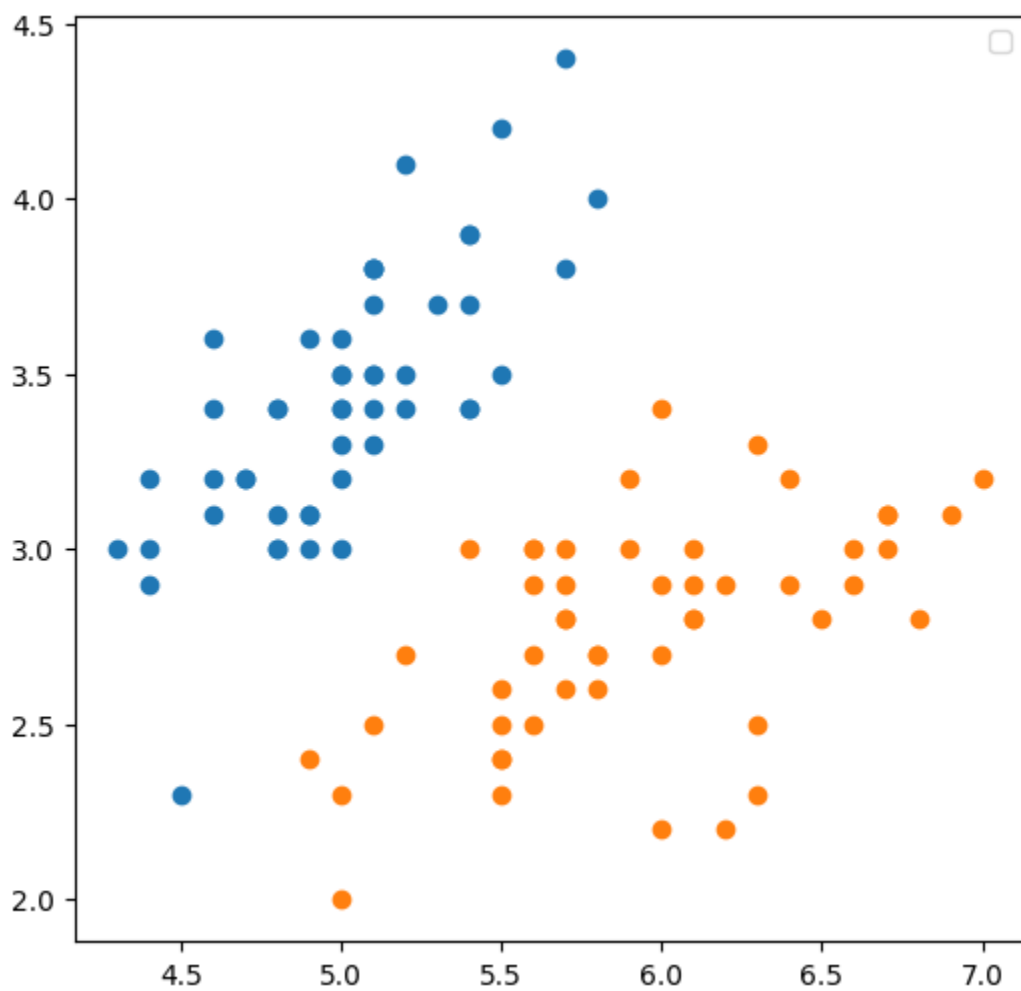
X: (100, 4)  
y: (100,)

Kết quả: (100, 4) (100,)

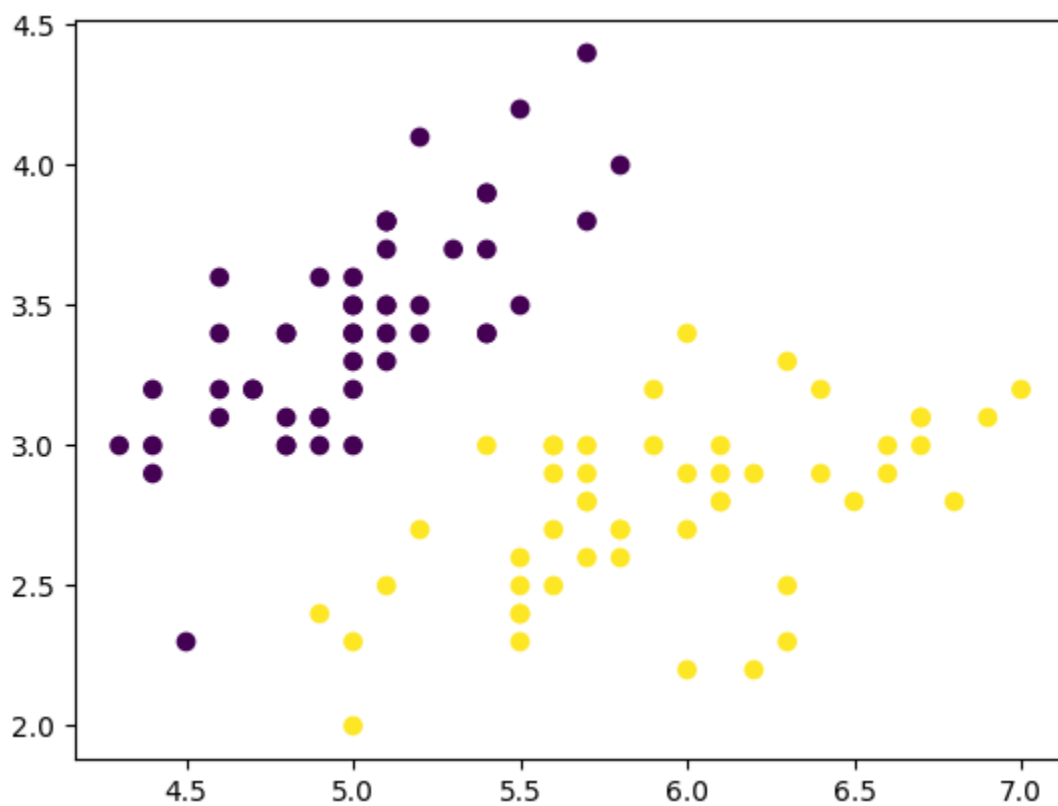
Kết quả: (100, 4) (100,)

```
In [148... # Trực quan hóa dữ liệu  
def visualize_data(X, y):  
    plt.figure(figsize=(13.5, 6))  
    plt.subplot(1, 2, 1)  
    plt.scatter(X[y == 0][:, 0], X[y == 0][:, 1],)  
    plt.scatter(X[y == 1][:, 0], X[y == 1][:, 1],)  
    plt.legend()  
    plt.show()  
visualize_data(X, y)
```

No artists with labels found to put in legend. Note that artists whose label start with an underscore are ignored when legend() is called with no argument.



Kết quả



```
In [148... # Bước 2: Định nghĩa mô hình hồi quy logistic bằng PyTorch
class LogisticRegressTorch(nn.Module):
    def __init__(self, n_features):
```

```
super(LogisticRegressTorch, self).__init__()
# Tạo một lớp tuyến tính (nn.Linear) với n_features đầu vào và 1 đầu ra
self.linear = nn.Linear(n_features, 1)
```

```
def forward(self, x):
    return torch.sigmoid(self.linear(x))
```

```
LogisticRegressTorch(n_features=4)
```

```
Out[1482]: LogisticRegressTorch(
  (linear): Linear(in_features=4, out_features=1, bias=True)
)
```

```
In [148... # Bước 3: Định nghĩa lớp dữ liệu
class IrisTorch(Dataset):
    def __init__(self, X, y):
        self.X = torch.tensor(X, dtype=torch.float32)
        self.y = torch.tensor(y, dtype=torch.float32).unsqueeze(1)

    def __len__(self):
        return len(self.X) # Trả về số lượng mẫu trong tập dữ liệu

    def __getitem__(self, idx):
        return self.X[idx], self.y[idx] # Trả về một cặp đặc trưng và nhãn tương ứng v
```

```
In [148... # Tạo dữ liệu
dataset = IrisTorch(X, y)

print(f"Số lượng mẫu trong tập dữ liệu: {len(dataset)}")
print("Một mẫu dữ liệu:", dataset[0])

Số lượng mẫu trong tập dữ liệu: 100
Một mẫu dữ liệu: (tensor([5.1000, 3.5000, 1.4000, 0.2000]), tensor([0.]))
```

```
In [148... # Bước 4: Chia tập dữ liệu thành tập huấn luyện và tập kiểm tra bằng cách chia ngẫu nhiên
dataset_size = len(dataset)
train_size = int(0.7 * dataset_size) # 70%
test_size = dataset_size - train_size # Còn lại là 30%

train_dataset, test_dataset = random_split(dataset, [train_size, test_size])
```

```
In [148... # Tạo DataLoader
batch_size = 64
train_loader = DataLoader(train_dataset, batch_size=batch_size, shuffle=True)
test_loader = DataLoader(test_dataset, batch_size=batch_size, shuffle=False)
```

```
In [148... # Bước 5: Định nghĩa criterion và optimizer
n_features = X.shape[1]
model = LogisticRegressTorch(n_features=n_features)
criterion = nn.BCELoss()
optimizer = optim.SGD(model.parameters(), lr=0.05)
```

```
In [148... # Huấn luyện mô hình
n_epochs = 200
train_losses = []
test_losses = []
test accuracies = []

for epoch in range(n_epochs):
    model.train()
    train_loss = 0.0
    for inputs, targets in train_loader:
        optimizer.zero_grad()
        outputs = model(inputs)
```

```

        loss = criterion(outputs, targets)
        loss.backward()
        optimizer.step()
        train_loss += loss.item() * inputs.size(0)

train_loss /= len(train_loader.dataset)
train_losses.append(train_loss)

# Đánh giá trên tập kiểm tra
model.eval()
test_loss = 0.0
correct = 0
total = 0
with torch.no_grad():
    for inputs, targets in test_loader:
        outputs = model(inputs)
        loss = criterion(outputs, targets)
        test_loss += loss.item() * inputs.size(0)

        predicted = (outputs >= 0.5).float()
        total += targets.size(0)
        correct += (predicted == targets).sum().item()

test_loss /= len(test_loader.dataset)
test_losses.append(test_loss)

accuracy = correct / total
test accuracies.append(accuracy)

print(f'Epoch {epoch+1}/{n_epochs}, Train Loss: {train_loss:.4f}, Test Loss: {test_l

```

```

Epoch 1/200, Train Loss: 1.0758, Test Loss: 0.9427, Test Accuracy: 0.4333
Epoch 2/200, Train Loss: 0.8932, Test Loss: 0.7720, Test Accuracy: 0.3333
Epoch 3/200, Train Loss: 0.7889, Test Loss: 0.7081, Test Accuracy: 0.5667
Epoch 4/200, Train Loss: 0.7626, Test Loss: 0.7034, Test Accuracy: 0.4333
Epoch 5/200, Train Loss: 0.6853, Test Loss: 0.6187, Test Accuracy: 0.5667
Epoch 6/200, Train Loss: 0.6639, Test Loss: 0.6103, Test Accuracy: 0.5667
Epoch 7/200, Train Loss: 0.6751, Test Loss: 0.5510, Test Accuracy: 0.9000
Epoch 8/200, Train Loss: 0.5643, Test Loss: 0.5180, Test Accuracy: 0.5667
Epoch 9/200, Train Loss: 0.5477, Test Loss: 0.4916, Test Accuracy: 0.6333
Epoch 10/200, Train Loss: 0.5165, Test Loss: 0.4692, Test Accuracy: 1.0000
Epoch 11/200, Train Loss: 0.4706, Test Loss: 0.4490, Test Accuracy: 1.0000
Epoch 12/200, Train Loss: 0.4441, Test Loss: 0.4278, Test Accuracy: 1.0000
Epoch 13/200, Train Loss: 0.4213, Test Loss: 0.4668, Test Accuracy: 0.9667
Epoch 14/200, Train Loss: 0.4218, Test Loss: 0.3895, Test Accuracy: 1.0000
Epoch 15/200, Train Loss: 0.3800, Test Loss: 0.3715, Test Accuracy: 1.0000
Epoch 16/200, Train Loss: 0.3631, Test Loss: 0.3505, Test Accuracy: 1.0000
Epoch 17/200, Train Loss: 0.3526, Test Loss: 0.3391, Test Accuracy: 1.0000
Epoch 18/200, Train Loss: 0.3343, Test Loss: 0.3248, Test Accuracy: 1.0000
Epoch 19/200, Train Loss: 0.3350, Test Loss: 0.3109, Test Accuracy: 1.0000
Epoch 20/200, Train Loss: 0.3083, Test Loss: 0.3129, Test Accuracy: 1.0000
Epoch 21/200, Train Loss: 0.2928, Test Loss: 0.2894, Test Accuracy: 1.0000
Epoch 22/200, Train Loss: 0.2829, Test Loss: 0.2897, Test Accuracy: 1.0000
Epoch 23/200, Train Loss: 0.2709, Test Loss: 0.2773, Test Accuracy: 1.0000
Epoch 24/200, Train Loss: 0.2607, Test Loss: 0.2651, Test Accuracy: 1.0000
Epoch 25/200, Train Loss: 0.2517, Test Loss: 0.2566, Test Accuracy: 1.0000
Epoch 26/200, Train Loss: 0.2432, Test Loss: 0.2492, Test Accuracy: 1.0000
Epoch 27/200, Train Loss: 0.2356, Test Loss: 0.2365, Test Accuracy: 1.0000
Epoch 28/200, Train Loss: 0.2337, Test Loss: 0.2296, Test Accuracy: 1.0000
Epoch 29/200, Train Loss: 0.2272, Test Loss: 0.2249, Test Accuracy: 1.0000
Epoch 30/200, Train Loss: 0.2152, Test Loss: 0.2368, Test Accuracy: 1.0000
Epoch 31/200, Train Loss: 0.2119, Test Loss: 0.2335, Test Accuracy: 1.0000
Epoch 32/200, Train Loss: 0.2060, Test Loss: 0.2084, Test Accuracy: 1.0000
Epoch 33/200, Train Loss: 0.1962, Test Loss: 0.2074, Test Accuracy: 1.0000
Epoch 34/200, Train Loss: 0.1906, Test Loss: 0.2018, Test Accuracy: 1.0000
Epoch 35/200, Train Loss: 0.1856, Test Loss: 0.1913, Test Accuracy: 1.0000

```

Epoch 36/200,	Train Loss: 0.1821,	Test Loss: 0.2018,	Test Accuracy: 1.0000
Epoch 37/200,	Train Loss: 0.1782,	Test Loss: 0.1818,	Test Accuracy: 1.0000
Epoch 38/200,	Train Loss: 0.1731,	Test Loss: 0.1869,	Test Accuracy: 1.0000
Epoch 39/200,	Train Loss: 0.1678,	Test Loss: 0.1757,	Test Accuracy: 1.0000
Epoch 40/200,	Train Loss: 0.1637,	Test Loss: 0.1714,	Test Accuracy: 1.0000
Epoch 41/200,	Train Loss: 0.1600,	Test Loss: 0.1717,	Test Accuracy: 1.0000
Epoch 42/200,	Train Loss: 0.1560,	Test Loss: 0.1726,	Test Accuracy: 1.0000
Epoch 43/200,	Train Loss: 0.1535,	Test Loss: 0.1710,	Test Accuracy: 1.0000
Epoch 44/200,	Train Loss: 0.1505,	Test Loss: 0.1628,	Test Accuracy: 1.0000
Epoch 45/200,	Train Loss: 0.1462,	Test Loss: 0.1522,	Test Accuracy: 1.0000
Epoch 46/200,	Train Loss: 0.1460,	Test Loss: 0.1509,	Test Accuracy: 1.0000
Epoch 47/200,	Train Loss: 0.1406,	Test Loss: 0.1485,	Test Accuracy: 1.0000
Epoch 48/200,	Train Loss: 0.1380,	Test Loss: 0.1442,	Test Accuracy: 1.0000
Epoch 49/200,	Train Loss: 0.1374,	Test Loss: 0.1433,	Test Accuracy: 1.0000
Epoch 50/200,	Train Loss: 0.1327,	Test Loss: 0.1405,	Test Accuracy: 1.0000
Epoch 51/200,	Train Loss: 0.1304,	Test Loss: 0.1379,	Test Accuracy: 1.0000
Epoch 52/200,	Train Loss: 0.1281,	Test Loss: 0.1368,	Test Accuracy: 1.0000
Epoch 53/200,	Train Loss: 0.1252,	Test Loss: 0.1349,	Test Accuracy: 1.0000
Epoch 54/200,	Train Loss: 0.1229,	Test Loss: 0.1316,	Test Accuracy: 1.0000
Epoch 55/200,	Train Loss: 0.1211,	Test Loss: 0.1289,	Test Accuracy: 1.0000
Epoch 56/200,	Train Loss: 0.1193,	Test Loss: 0.1294,	Test Accuracy: 1.0000
Epoch 57/200,	Train Loss: 0.1166,	Test Loss: 0.1270,	Test Accuracy: 1.0000
Epoch 58/200,	Train Loss: 0.1146,	Test Loss: 0.1291,	Test Accuracy: 1.0000
Epoch 59/200,	Train Loss: 0.1131,	Test Loss: 0.1299,	Test Accuracy: 1.0000
Epoch 60/200,	Train Loss: 0.1125,	Test Loss: 0.1384,	Test Accuracy: 1.0000
Epoch 61/200,	Train Loss: 0.1140,	Test Loss: 0.1198,	Test Accuracy: 1.0000
Epoch 62/200,	Train Loss: 0.1070,	Test Loss: 0.1175,	Test Accuracy: 1.0000
Epoch 63/200,	Train Loss: 0.1056,	Test Loss: 0.1135,	Test Accuracy: 1.0000
Epoch 64/200,	Train Loss: 0.1049,	Test Loss: 0.1114,	Test Accuracy: 1.0000
Epoch 65/200,	Train Loss: 0.1038,	Test Loss: 0.1131,	Test Accuracy: 1.0000
Epoch 66/200,	Train Loss: 0.1008,	Test Loss: 0.1127,	Test Accuracy: 1.0000
Epoch 67/200,	Train Loss: 0.0994,	Test Loss: 0.1114,	Test Accuracy: 1.0000
Epoch 68/200,	Train Loss: 0.0980,	Test Loss: 0.1063,	Test Accuracy: 1.0000
Epoch 69/200,	Train Loss: 0.0971,	Test Loss: 0.1058,	Test Accuracy: 1.0000
Epoch 70/200,	Train Loss: 0.0954,	Test Loss: 0.1035,	Test Accuracy: 1.0000
Epoch 71/200,	Train Loss: 0.0944,	Test Loss: 0.1027,	Test Accuracy: 1.0000
Epoch 72/200,	Train Loss: 0.0930,	Test Loss: 0.1007,	Test Accuracy: 1.0000
Epoch 73/200,	Train Loss: 0.0922,	Test Loss: 0.1009,	Test Accuracy: 1.0000
Epoch 74/200,	Train Loss: 0.0904,	Test Loss: 0.0981,	Test Accuracy: 1.0000
Epoch 75/200,	Train Loss: 0.0899,	Test Loss: 0.0968,	Test Accuracy: 1.0000
Epoch 76/200,	Train Loss: 0.0888,	Test Loss: 0.0962,	Test Accuracy: 1.0000
Epoch 77/200,	Train Loss: 0.0873,	Test Loss: 0.0960,	Test Accuracy: 1.0000
Epoch 78/200,	Train Loss: 0.0860,	Test Loss: 0.0954,	Test Accuracy: 1.0000
Epoch 79/200,	Train Loss: 0.0849,	Test Loss: 0.0964,	Test Accuracy: 1.0000
Epoch 80/200,	Train Loss: 0.0838,	Test Loss: 0.0947,	Test Accuracy: 1.0000
Epoch 81/200,	Train Loss: 0.0828,	Test Loss: 0.0969,	Test Accuracy: 1.0000
Epoch 82/200,	Train Loss: 0.0823,	Test Loss: 0.0934,	Test Accuracy: 1.0000
Epoch 83/200,	Train Loss: 0.0809,	Test Loss: 0.0915,	Test Accuracy: 1.0000
Epoch 84/200,	Train Loss: 0.0799,	Test Loss: 0.0908,	Test Accuracy: 1.0000
Epoch 85/200,	Train Loss: 0.0791,	Test Loss: 0.0922,	Test Accuracy: 1.0000
Epoch 86/200,	Train Loss: 0.0787,	Test Loss: 0.0935,	Test Accuracy: 1.0000
Epoch 87/200,	Train Loss: 0.0783,	Test Loss: 0.0889,	Test Accuracy: 1.0000
Epoch 88/200,	Train Loss: 0.0766,	Test Loss: 0.0884,	Test Accuracy: 1.0000
Epoch 89/200,	Train Loss: 0.0758,	Test Loss: 0.0876,	Test Accuracy: 1.0000
Epoch 90/200,	Train Loss: 0.0750,	Test Loss: 0.0861,	Test Accuracy: 1.0000
Epoch 91/200,	Train Loss: 0.0740,	Test Loss: 0.0849,	Test Accuracy: 1.0000
Epoch 92/200,	Train Loss: 0.0733,	Test Loss: 0.0840,	Test Accuracy: 1.0000
Epoch 93/200,	Train Loss: 0.0725,	Test Loss: 0.0816,	Test Accuracy: 1.0000
Epoch 94/200,	Train Loss: 0.0717,	Test Loss: 0.0819,	Test Accuracy: 1.0000
Epoch 95/200,	Train Loss: 0.0709,	Test Loss: 0.0798,	Test Accuracy: 1.0000
Epoch 96/200,	Train Loss: 0.0703,	Test Loss: 0.0796,	Test Accuracy: 1.0000
Epoch 97/200,	Train Loss: 0.0695,	Test Loss: 0.0778,	Test Accuracy: 1.0000
Epoch 98/200,	Train Loss: 0.0690,	Test Loss: 0.0775,	Test Accuracy: 1.0000
Epoch 99/200,	Train Loss: 0.0682,	Test Loss: 0.0778,	Test Accuracy: 1.0000
Epoch 100/200,	Train Loss: 0.0675,	Test Loss: 0.0761,	Test Accuracy: 1.0000
Epoch 101/200,	Train Loss: 0.0668,	Test Loss: 0.0763,	Test Accuracy: 1.0000

[illegible]

```

Epoch 168/200, Train Loss: 0.0409, Test Loss: 0.0470, Test Accuracy: 1.0000
Epoch 169/200, Train Loss: 0.0406, Test Loss: 0.0470, Test Accuracy: 1.0000
Epoch 170/200, Train Loss: 0.0403, Test Loss: 0.0467, Test Accuracy: 1.0000
Epoch 171/200, Train Loss: 0.0402, Test Loss: 0.0463, Test Accuracy: 1.0000
Epoch 172/200, Train Loss: 0.0400, Test Loss: 0.0461, Test Accuracy: 1.0000
Epoch 173/200, Train Loss: 0.0398, Test Loss: 0.0453, Test Accuracy: 1.0000
Epoch 174/200, Train Loss: 0.0397, Test Loss: 0.0451, Test Accuracy: 1.0000
Epoch 175/200, Train Loss: 0.0394, Test Loss: 0.0454, Test Accuracy: 1.0000
Epoch 176/200, Train Loss: 0.0391, Test Loss: 0.0449, Test Accuracy: 1.0000
Epoch 177/200, Train Loss: 0.0389, Test Loss: 0.0447, Test Accuracy: 1.0000
Epoch 178/200, Train Loss: 0.0387, Test Loss: 0.0446, Test Accuracy: 1.0000
Epoch 179/200, Train Loss: 0.0384, Test Loss: 0.0444, Test Accuracy: 1.0000
Epoch 180/200, Train Loss: 0.0382, Test Loss: 0.0445, Test Accuracy: 1.0000
Epoch 181/200, Train Loss: 0.0380, Test Loss: 0.0439, Test Accuracy: 1.0000
Epoch 182/200, Train Loss: 0.0379, Test Loss: 0.0440, Test Accuracy: 1.0000
Epoch 183/200, Train Loss: 0.0376, Test Loss: 0.0450, Test Accuracy: 1.0000
Epoch 184/200, Train Loss: 0.0375, Test Loss: 0.0438, Test Accuracy: 1.0000
Epoch 185/200, Train Loss: 0.0372, Test Loss: 0.0432, Test Accuracy: 1.0000
Epoch 186/200, Train Loss: 0.0371, Test Loss: 0.0422, Test Accuracy: 1.0000
Epoch 187/200, Train Loss: 0.0371, Test Loss: 0.0422, Test Accuracy: 1.0000
Epoch 188/200, Train Loss: 0.0368, Test Loss: 0.0420, Test Accuracy: 1.0000
Epoch 189/200, Train Loss: 0.0366, Test Loss: 0.0424, Test Accuracy: 1.0000
Epoch 190/200, Train Loss: 0.0363, Test Loss: 0.0421, Test Accuracy: 1.0000
Epoch 191/200, Train Loss: 0.0362, Test Loss: 0.0415, Test Accuracy: 1.0000
Epoch 192/200, Train Loss: 0.0361, Test Loss: 0.0416, Test Accuracy: 1.0000
Epoch 193/200, Train Loss: 0.0358, Test Loss: 0.0419, Test Accuracy: 1.0000
Epoch 194/200, Train Loss: 0.0356, Test Loss: 0.0414, Test Accuracy: 1.0000
Epoch 195/200, Train Loss: 0.0354, Test Loss: 0.0413, Test Accuracy: 1.0000
Epoch 196/200, Train Loss: 0.0352, Test Loss: 0.0412, Test Accuracy: 1.0000
Epoch 197/200, Train Loss: 0.0351, Test Loss: 0.0403, Test Accuracy: 1.0000
Epoch 198/200, Train Loss: 0.0350, Test Loss: 0.0403, Test Accuracy: 1.0000
Epoch 199/200, Train Loss: 0.0348, Test Loss: 0.0405, Test Accuracy: 1.0000
Epoch 200/200, Train Loss: 0.0346, Test Loss: 0.0410, Test Accuracy: 1.0000

```

In [148..

```
# Vẽ đồ thị Train Loss và Test Loss
```

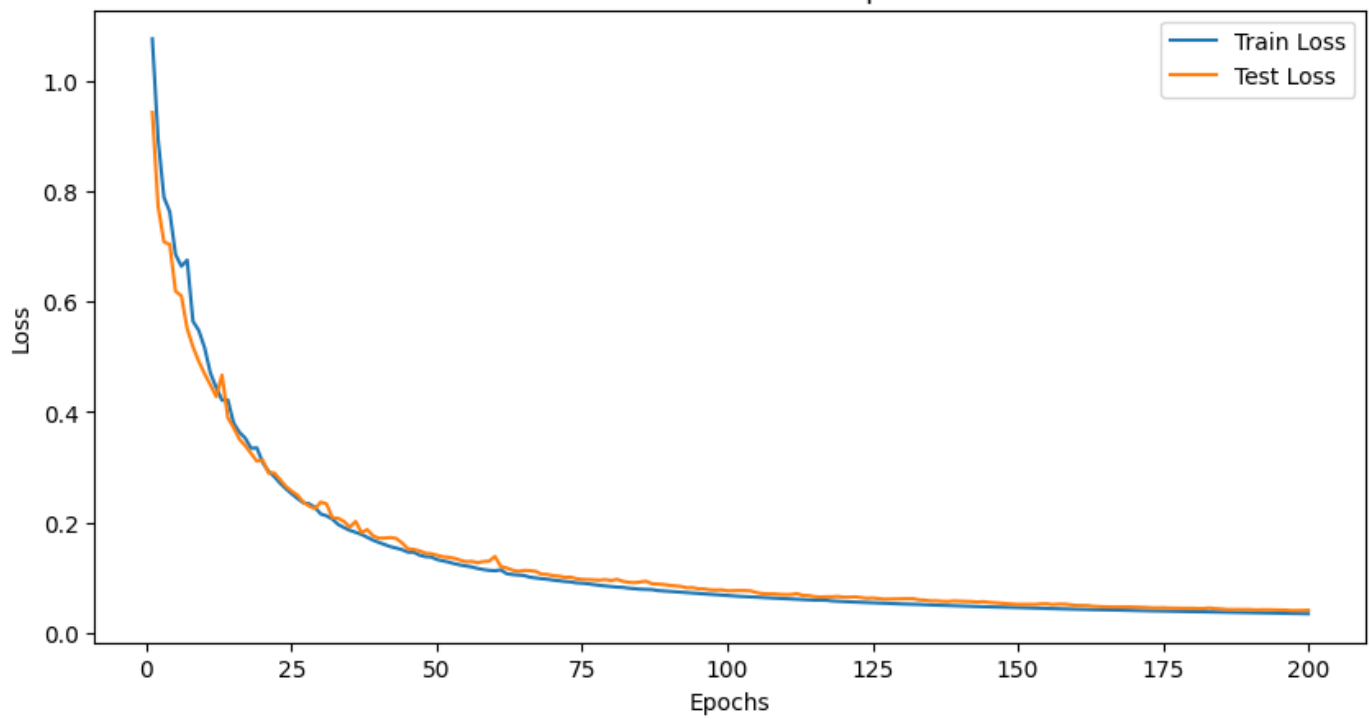
```
plt.figure(figsize=(10, 5))
plt.plot(range(1, n_epochs+1), train_losses, label='Train Loss')
plt.plot(range(1, n_epochs+1), test_losses, label='Test Loss')
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.legend()
plt.title('Train and Test Loss over Epochs')
plt.show()
```

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

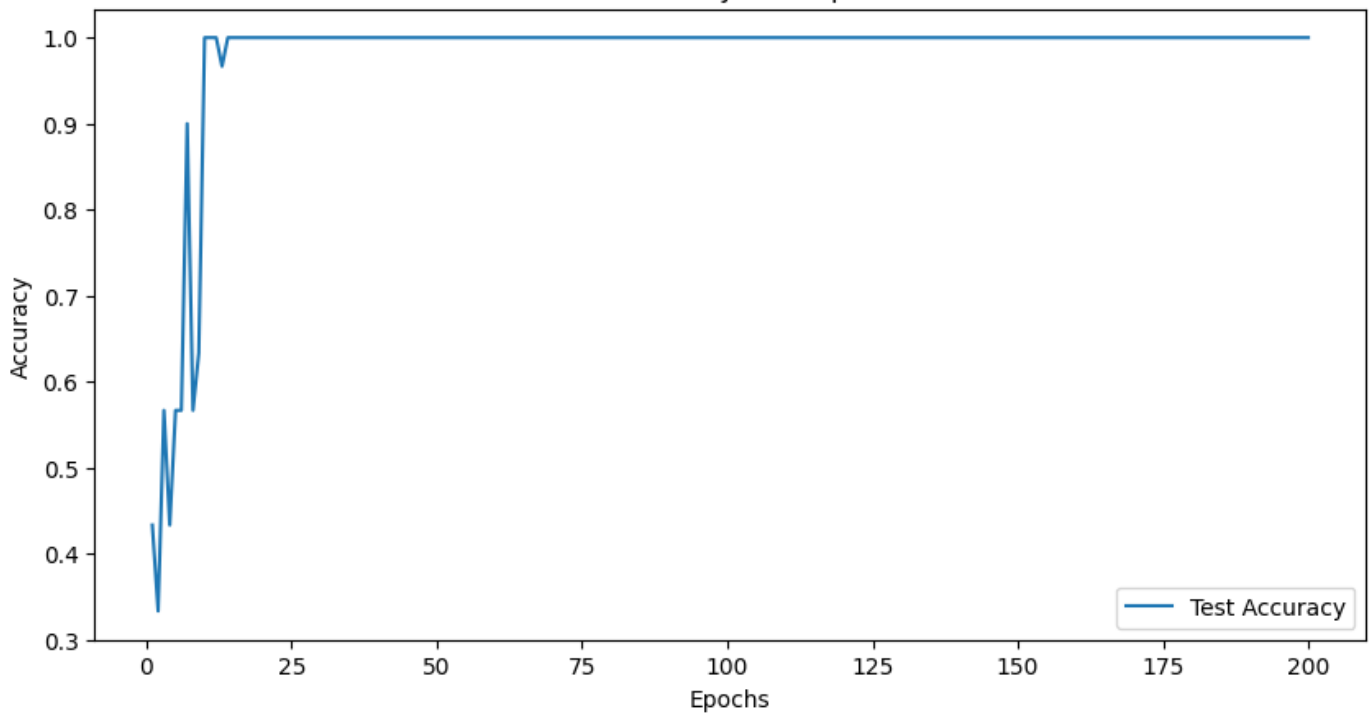
```
# Vẽ đồ thị Test Accuracy
```

```
plt.figure(figsize=(10, 5))
plt.plot(range(1, n_epochs+1), test accuracies, label='Test Accuracy')
plt.xlabel('Epochs')
plt.ylabel('Accuracy')
plt.legend()
plt.title('Test Accuracy over Epochs')
plt.show()
```

Train and Test Loss over Epochs



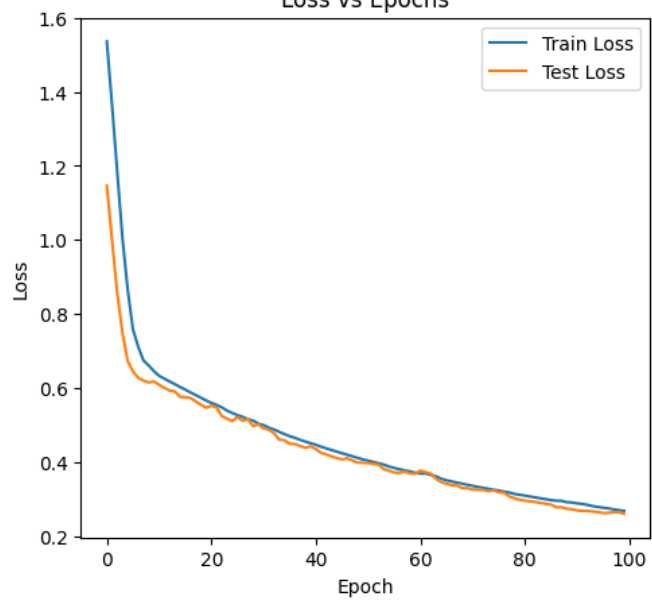
Test Accuracy over Epochs



Kết quả:



Loss vs Epochs



Accuracy vs Epochs

