

# Báo cáo bài tập thực hành 01

## 1. Đề bài.

Xây dựng mô hình mạng DNN từ đầu (from scratch), không sử dụng các thư viện deep learning có sẵn như TensorFlow, PyTorch để phân loại ảnh chữ số viết tay MNIST.

## 2. Mô hình.

Mô hình DNN bao gồm:

- Một lớp đầu vào có kích thước 784 (do dữ liệu MNIST có kích thước ảnh 28x28).
- Một lớp ẩn với kích thước được thay đổi tùy theo từng thử nghiệm.
- Một lớp đầu ra với 10 đơn vị, mỗi đơn vị tương ứng với một chữ số từ 0 đến 9.
- Các hàm kích hoạt khác nhau cho lớp ẩn: relu, sigmoid, tanh, leaky\_relu, và softmax cho lớp đầu ra.
- Quá trình huấn luyện sử dụng gradient descent với các tham số như tốc độ học (learning\_rate), kích thước batch (batch\_size), và số epoch.

## 3. Siêu tham số.

### 1. ReLU (Rectified Linear Unit)

- Batch Size: 32
- Learning Rate: 0.1
- Hidden Layer Size: 16

### 2. Sigmoid

- Batch Size: 16
- Learning Rate: 0.1
- Hidden Layer Size: 64

### 3. Tanh (Hyperbolic Tangent)

- Batch Size: 64
- Learning Rate: 0.1
- Hidden Layer Size: 32

#### 4. Leaky ReLU

- Batch Size: 128
- Learning Rate: 0.1
- Hidden Layer Size: 128

#### 5. Softmax (chỉ ở lớp đầu ra)

- Batch Size: 32
- Learning Rate: 0.1
- Hidden Layer Size: 128

### 4. Kết quả.

Các kết quả huấn luyện của bạn cho các hàm kích hoạt khác nhau là:

Hàm kích hoạt	Batch size	Hidden size	Test Accuracy (Mean $\pm$ Std)
ReLU 32	32	16	0.9518 $\pm$ 0.0018
Sigmoid	16	64	0.9695 $\pm$ 0.0017
Tanh	64	32	0.9686 $\pm$ 0.0020
Leaky ReLU	128	128	0.9683 $\pm$ 0.0016
Softmax	32	128	0.9222 $\pm$ 0.0013

### 5. Nhận xét.

Qua quá trình thử nghiệm, em thấy sigmoid cho kết quả cao nhất và khá ổn định, rất phù hợp với bài toán đơn giản như MNIST. ReLU cũng hoạt động tốt, độ lệch chuẩn thấp nên mô hình học ổn định. Tanh cho kết quả ở mức khá nhưng không nổi bật. Leaky ReLU thì hoạt động ổn, nhưng không vượt trội. Riêng softmax dùng trong lớp ẩn không phù hợp nên kết quả thấp nhất là điều dễ hiểu.

### 6. Kết luận.

Dựa trên kết quả thực nghiệm, em thấy Sigmoid và ReLU là hai hàm kích hoạt hoạt động tốt nhất với mô hình DNN đơn giản trên MNIST. Mô hình cho kết quả ổn định với độ chính xác luôn trên 95%.