Họ và tên: Trần Quang Dân

MSSV: 21521917

Lớp: DS201.O11.2

**BÀI THỰC HÀNH 2. TỐI ƯU MÔ HÌNH MẠNG NEURAL**

**A. KHỞI ĐỘNG**

* Một số tham số (parameter) của mô hình mạng neural?

Weights: Đây là các tham số cần được học trong quá trình huấn luyện mô hình để điều chỉnh đầu ra của các lớp.

Bias : Bias được sử dụng để điều chỉnh đầu ra của các neuron trong mạng.

* Một số siêu tham số (hyperparameter) của mô hình mạng neural?

Hidden layers và Hidden size: Xác định kiến trúc tổng quan của mạng neural.

Learning rate: Điều chỉnh tốc độ học của mô hình trong quá trình huấn luyện.

Epochs: Số lần huấn luyện sẽ được thực hiện trên toàn bộ tập dữ liệu.

Batch size: Số lượng điểm dữ liệu sẽ được xử lý trong mỗi lần cập nhật tham số trong quá trình huấn luyện.

* Một số phương pháp tối ưu mô hình mạng neural đã học?

Gradient Descent (GD): Phương pháp cơ bản để tối ưu hóa mô hình bằng cách điều chỉnh tham số dựa trên gradient của hàm mất mát.

Stochastic Gradient Descent (SGD): Một biến thể của Gradient Descent sử dụng từng mẫu dữ liệu một để cập nhật tham số.

Adam: Một thuật toán tối ưu hóa tỷ lệ học tập thích nghi và có hiệu suất tốt trong việc huấn luyện mạng neural.

RMSProp: Một phương pháp tối ưu hóa sử dụng giá trị gradient chuẩn hóa để điều chỉnh hệ số học.

**B. BỘ DỮ LIỆU**

1. Load bộ dữ liệu Fashion MNIST?

A screenshot of a computer program

Description automatically generated

1. Khảo sát bộ dữ liệu Fashion MNIST

* Tập train có tổng cộng 10 nhãn, gồm: 'T-shirt/top', 'Trouser', 'Pullover', 'Dress', 'Coat', 'Sandal', 'Shirt', 'Sneaker', 'Bag', 'Ankle boot'.
* Tập train có 60000 ảnh, tập test có 10000 ảnh.
* Mỗi ảnh trong tập train và tập test có kích thước 28x28

1. Trực quan hóa các ảnh sau: train[0], train[50], test[100], test[1000]?

A collage of images of clothes

Description automatically generated

**C. XÂY DỰNG VÀ HUẤN LUYỆN MÔ HÌNH - HÀM KÍCH HOẠT**

1. Kể tên một số hàm kích hoạt đã học?

* Hàm Sigmoid: Hàm này chuyển đổi đầu vào thành một phạm vi giữa 0 và 1. Thường được sử dụng trong các lớp ẩn của mạng neural nhỏ hoặc trong bài toán phân loại nhị phân.
* Hàm ReLU: Hàm này ánh xạ tất cả các giá trị âm thành 0 và giữ nguyên các giá trị dương. Đây là một trong những hàm kích hoạt phổ biến nhất trong mạng neural hiện đại.
* Hàm Tanh: Hàm này chuyển đổi đầu vào thành một phạm vi giữa -1 và 1. Nó tương tự như hàm sigmoid, nhưng có giá trị trung bình là 0.
* Hàm Softmax: Thường được sử dụng trong lớp đầu ra của mạng neural để biến đầu ra thành một phân phối xác suất trên nhiều lớp. Thường được sử dụng trong bài toán phân loại nhiều lớp.

1. Xây dựng mạng neural đơn giản có cấu trúc như hình dưới đây?

A screen shot of a computer program

Description automatically generated

1. Xem cấu trúc của mô hình đã xây dựng và cho biết tổng số tham số?

A screenshot of a computer program

Description automatically generated

Tổng số tham số: 535.818 tham số

1. Huấn luyện mô hình với các thông số sau:

A black screen with white text

Description automatically generated

1. Đánh giá mô hình trên tập test và thực hiện các yêu cầu sau:

A graph with numbers and symbols

Description automatically generated with medium confidence

A screenshot of a computer screen

Description automatically generated

**D. HÀM MẤT MÁT**

Kể tên một số hàm mất mát đã học?

* Hàm Cross-Entropy (Categorical Cross-Entropy): Thường được sử dụng trong bài toán phân loại nhiều lớp, đo lường sự sai khác giữa phân phối xác suất dự đoán và phân phối xác suất thực tế.
* Hàm Mean Squared Error (MSE): Thường được sử dụng trong bài toán hồi quy, đo lường sự sai khác giữa các dự đoán và giá trị thực tế dưới dạng bình phương.
* Hàm Binary Cross-Entropy (BCE): Thường được sử dụng trong bài toán phân loại nhị phân, đo lường sự sai khác giữa phân phối xác suất dự đoán và phân phối xác suất thực tế.

E. CHUẨN HÓA MÔ HÌNH

1. Viết công thức chuẩn hóa L1 và L2?

A black and white text

Description automatically generated

L2:

A black text on a white background

Description automatically generated

L1:

1. Xây dựng lại mô hình tương tự như mục C và chuẩn hoá L2 cho các tham số của mô hình với hệ số lambda là 0.01

A graph of numbers and symbols

Description automatically generated with medium confidence

A screenshot of a computer screen

Description automatically generated

**F. KHỞI TẠO THAM SỐ**

1. Kể tên một số phương pháp khởi tạo tham số đã học?

* Khởi tạo ngẫu nhiên (Random Initialization): Tham số được khởi tạo ngẫu nhiên từ một phân phối đều trong một khoảng nhất định.
* Khởi tạo zeros (Zero Initialization): Tất cả tham số được khởi tạo bằng giá trị 0. Tuy nhiên, đây không phải là một phương pháp khởi tạo tham số phổ biến do nó có thể dẫn đến các vấn đề đối với mạng neural.
* Khởi tạo He (He Initialization): Được thiết kế đặc biệt cho các hàm kích hoạt ReLU và biến thể của nó. Tham số được khởi tạo từ một phân phối cân bằng.
* Khởi tạo Xavier/Glorot (Xavier Initialization): Cũng dành riêng cho các hàm kích hoạt ReLU và biến thể của nó. Tham số được khởi tạo từ một phân phối cân bằng, được điều chỉnh dựa trên kích thước đầu vào và đầu ra của lớp.

1. Xây dựng lại mô hình tương tự như mục C với các tham số W và b được khởi tạo theo phương pháp Ones

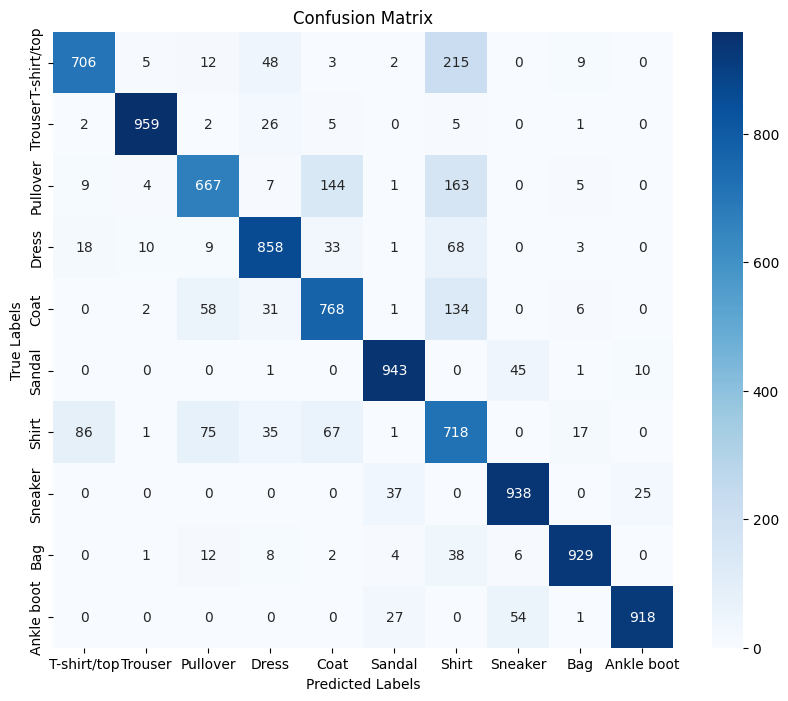
A diagram of a number of clothing

Description automatically generated with medium confidence

A screenshot of a computer screen

Description automatically generated

1. Xây dựng lại mô hình tương tự như mục C với các tham số W được khởi tạo theo phương pháp Glorot Uniform và các tham số b được khởi tạo theo phương pháp Zeros.



A screenshot of a computer screen

Description automatically generated

**G. CÁC THUẬT TOÁN TỐI ƯU**

1. Kể tên một số thuật toán tối ưu đã học?

* Stochastic Gradient Descent (SGD): Là một phương pháp tối ưu cơ bản, cập nhật trọng số theo hướng ngược của độ dốc.
* Adam (Adaptive Moment Estimation): Kết hợp cả động độ của RMSprop và động độ của Momentum. Thường được sử dụng trong nhiều ứng dụng vì khả năng tự điều chỉnh tốc độ học.
* RMSprop (Root Mean Square Propagation): Cập nhật trọng số dựa trên tỷ lệ giảm độ dốc của mỗi tham số.

1. Xây dựng lại mô hình tương tự như mục C và thực hiện các yêu cầu sau:

* RMSprop

A screen shot of a computer

Description automatically generated A graph of different sizes and colors

Description automatically generated with medium confidence

* Adam

A graph with numbers and symbols

Description automatically generated with medium confidence

A screenshot of a computer screen

Description automatically generated