**Câu 1: Giải thích steps\_per\_epoch, epoch?**

Bài làm:

Một epoch được định nghĩa là một lần mô hình mạng quét toàn bộ dữ liệu huấn luyện. Khái niệm thường hay bị nhầm lẫn với iteration. Mỗi iteration được xem là một lần forward pass (tiến) và backward pass (lùi) của một batch.

Ví dụ, một dataset có 20 mẫu dữ liệu, ta đặt kích thước batch là 2, epoch là 3.

Vậy trong mỗi epoch, ta có tổng cộng 10 batch (20/2). Mỗi batch sẽ đưa vào mô hình và thực hiện thuật toán lùi 10 lần, ứng với 10 iteration mỗi epoch. Tổng cộng có 10 \* 3 == 30 iteration trong cả 3 epoch.

step\_per\_epoch là số batch trong mỗi epoch. Ở ví dụ trên, nó chính là tổng số batch (10) đã tính ở trên.

**Câu 2: Mô tả code: Mô hình gồm bao nhiêu lớp? Đầu vào và đầu ra là gì? Bao nhiêu ảnh train bao nhiêu ảnh test? Nguồn ảnh? kích thước ảnh ban đầu?**

Bài làm:

Nếu tính cả lớp vào và ra (input, output layer), ta có 6 lớp. Nếu chỉ xét lớp ẩn, ta có 4 lớp tất cả. Dữ liệu đầu vào là ma trận ba chiều, cấu thành từ các ma trận hai chiều đại diện cho các ảnh. Đầu ra là ma trận hai chiều.

Về nguồn, ảnh lấy từ <http://yann.lecun.com/exdb/mnist/> (kiểu ubyte, gốc) hoặc <https://pjreddie.com/projects/mnist-in-csv/> (kiểu csv).

Có 60.000 ảnh được train, 10.000 ảnh được test (theo bộ dataset MNIST). Mỗi ảnh có kích thước ban đầu là 24x24 (2D).