

Personal Note 07 - Sequential Model

1. Sequential Model

- **Định nghĩa:** Là một loại mô hình trong Keras (thư viện deep learning) cho phép xây dựng mạng nơ-ron bằng cách xếp chồng các lớp theo thứ tự tuần tự.
 - **Cách hoạt động:** Mỗi lớp trong mô hình nhận đầu ra từ lớp trước đó và truyền nó cho lớp tiếp theo.
- **Dễ dàng xây dựng:** Thích hợp cho các mô hình đơn giản, nơi dữ liệu đi qua từng lớp một.
- **Thích hợp cho các loại mạng nơ-ron:** Đặc biệt cho mạng nơ-ron tích chập (CNN) và mạng nơ-ron hồi tiếp (RNN).

2. LSTM Model

- **Định nghĩa:** LSTM là một loại mạng nơ-ron hồi tiếp (RNN) được thiết kế để học các phụ thuộc dài hạn trong chuỗi dữ liệu.
- **Chức năng:** Giúp xử lý và dự đoán chuỗi dữ liệu với thông tin được lưu trữ qua nhiều bước thời gian.
- **Cấu trúc :** LSTM bao gồm các đơn vị nhớ (memory cell) với ba cổng:
 - **Cổng vào (Input Gate):** Quyết định thông tin nào sẽ được thêm vào đơn vị nhớ.
 - **Cổng quên (Forget Gate):** Quyết định thông tin nào sẽ được loại bỏ khỏi đơn vị nhớ.
 - **Cổng đầu ra (Output Gate):** Quyết định thông tin nào sẽ được truyền ra khỏi đơn vị nhớ.
- **Cách hoạt động:**

- **Xử lý chuỗi:** Mỗi bước thời gian, LSTM nhận đầu vào và trạng thái ẩn từ bước trước đó, sau đó cập nhật trạng thái nhớ dựa trên cổng.
- **Lưu trữ thông tin:** LSTM có khả năng giữ lại thông tin trong thời gian dài, vượt qua hạn chế của RNN thông thường.

3. Text Generation problem

- Là thao tác tự động tạo ra văn bản giống như văn bản con người viết, dựa trên một đầu vào hoặc gợi ý.

4. LSTM model for text generation problem

- B1 (quan trọng nhất): Chuẩn bị dữ liệu
 - Tiền xử lý: tách thành các tokens → chuyển đổi thành dạng số
- B2: Xây dựng mô hình LSTM
- B3: Train & Test