

Dự án cuối kỳ
Môn: Nhập môn học máy

Cho dữ liệu thuộc dạng Time Series.
Sử dụng dữ liệu trên trang UCI hoặc Kaggle

<https://archive.ics.uci.edu/ml/datasets.php>

<https://www.kaggle.com/datasets>

Câu 1: (3.5 điểm)

Viết chương trình dự đoán cho bài toán Time Series sử dụng mạng Feed Forward Neural Networks (Perceptron đa tầng).

Câu 2: (3.5 điểm)

Vẫn sử dụng dữ liệu của bài toán 1. Viết chương trình sử dụng mạng RNN (hoặc các mạng cải tiến của nó như LSTM, GRU) để giải quyết bài toán.

Đối với câu 1 và câu 2:

- Chạy chương trình với nhiều cấu hình khác nhau: số lượng tầng ẩn, số nút tại mỗi tầng ẩn; hàm kích hoạt.
- Số lượng epoch khác nhau, batch size khác nhau
- Sử dụng early stopping để giải quyết hiện tượng overfitting. Vẽ đồ thị thể hiện accuracy của tập train và tập validation.

Câu 3 (3 điểm): trình bày lý thuyết mạng Recurrent Neural Network.

Lưu ý:

- 1) Hai câu 1 và 2 trong cùng 1 chương trình
- 2) Lý thuyết: trình bày mạng RNN và khuyến khích thêm LSTM