**Mục tiêu của mô hình :**

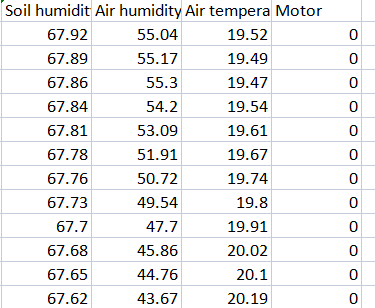
Hiện nay ,việc tới nước cho cây trồng ở nước vẫn còn công việc cuả con người và hoàn toàn là thủ công ,do đó ta cần một giải pháp để có thể tự đông hóa việc tưới nước .Nhận ra được điều đó, nhóm đã xây dựng một mô hình để thực hiện việc tưới nước tự động.Mô hình sẽ dựa vào các lần gửi dữ liệu trước đó để đưa ra quyết rằng trong 5 phút tiếp theo motor nước có được bật hay không.

**Giới thiệu data :**

Link dữ liệu gốc :<https://github.com/ayamlearning/Wazihub_Soil_Moisture_Prediction/blob/master/Dataset/Train.csv>

Dữ liệu được lấy cuộc thi “Wazihub Soil Moisture Prediction Challenge” - https://zindi.africa/competitions/wazihub-soil-moisture-prediction-challenge .Mỗi lần lần nhận dữ liệu từ sensor cách nhau 5 phút

Nhóm đã tiến hành trích xuất và cắt lọc để có được dư liệu phù hợp với mô hình và mục tiêu nhóm đã đặt ra ,dữ liệu ra khi xử lý sẽ có dạng :



Trong đó :

Soil humidity : độ ẩm đất

Air humidity (%) : độ ẩm không khí

Air temperature (C):nhiệt độ không khi

Motor: bật /tắt motor nước

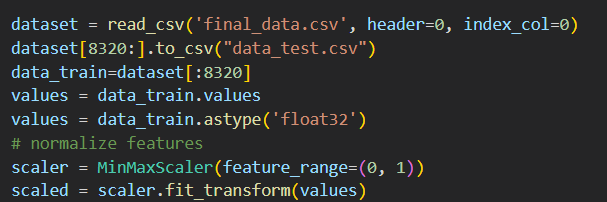
**Hiện thực :**

Vì dữ liệu của bài toán mang tính tuần tự và có tính chất Time Series nên mô hình LSTM rất thích hợp để sử dụng bởi LSTM vì LSTM có thể đưa ra output dựa trên những dữ liệu trước đó :

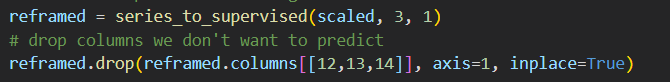
Đầu tiên thực hiện load dữ liệu :

Dữ liệu chính được chứa trong file **final\_data.csv** .Nhóm đã quyết định chia dữ liệu ra thành 2 phần : đó là **data\_train** dùng để train model và **data\_test** được lưu trong **data\_test.csv** dùng để test model.

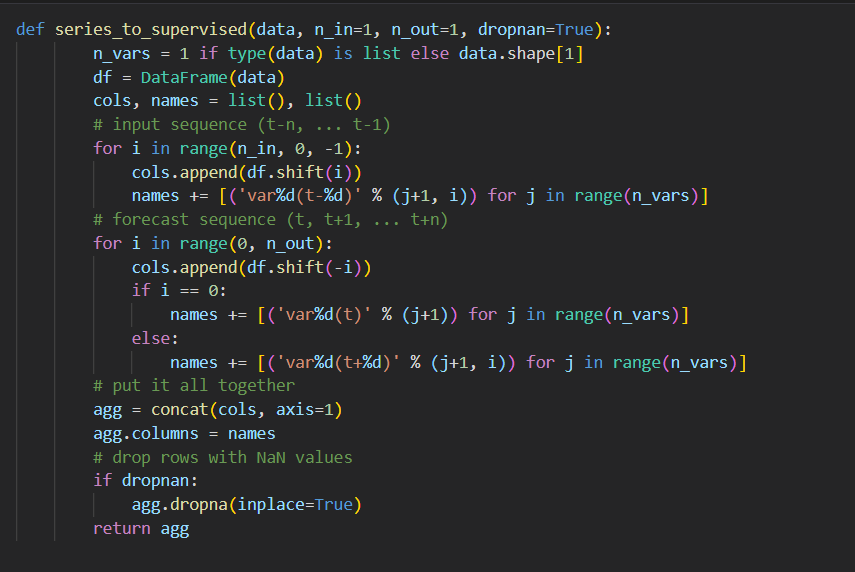
Trước khi đưa dữ liệu vào để huấn luyện mô hình, nhóm chuẩn hoá lại dữ liệu sao cho khoảng giá trị của dữ liệu sẽ nằm giữa 0 và 1.



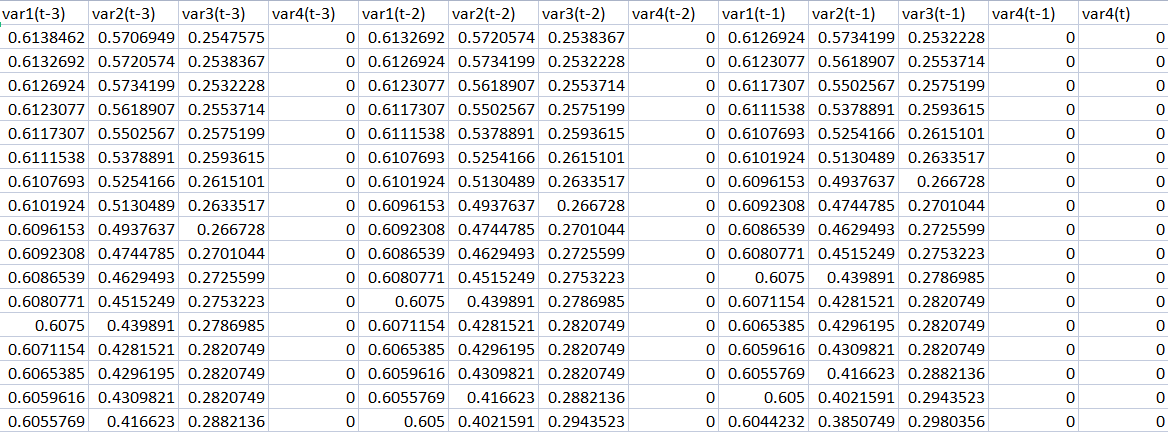
Sau khi chuẩn hóa dữ liệu  tập dữ liệu được chuyển thành dữ liệu học tập có giám sát để có model có thể học được



Hàm chuyển dữ liệu thành dữ liệu học tập có giám sát:

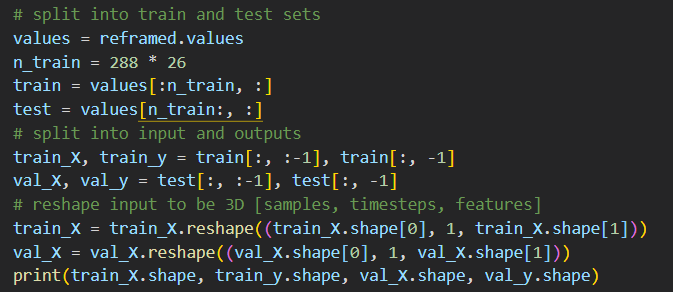


Dữ liệu sau khi được chuyển về dữ liệu học tập có giám sát sẽ có dạng :



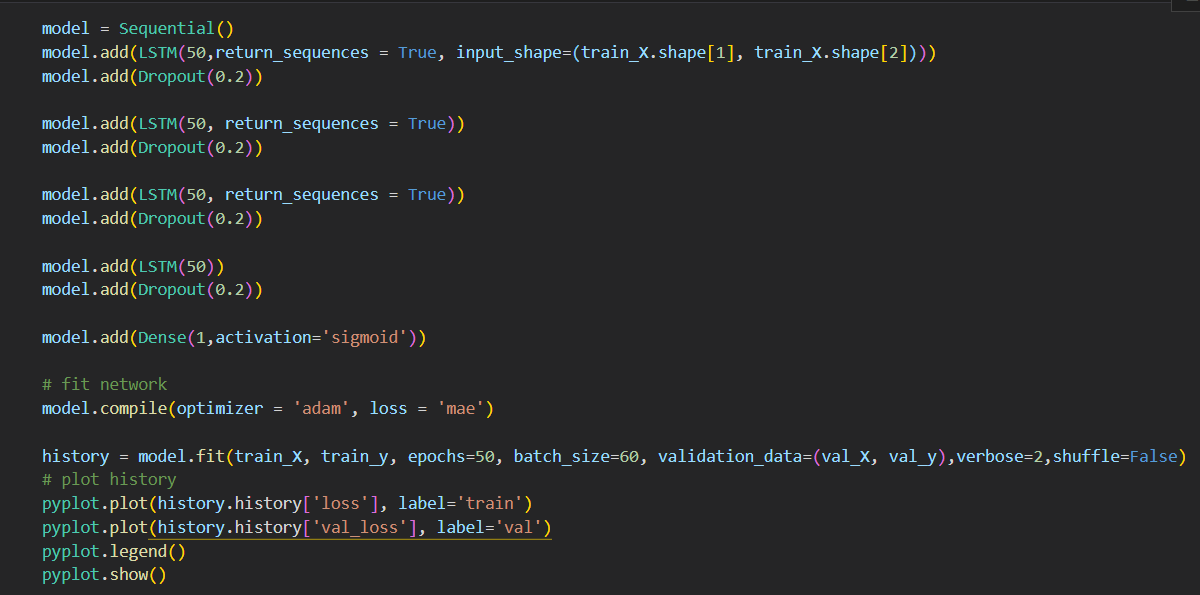
Trong đó n lần quan sát gần nhất sẽ được sử dụng để đưa ra dựa đoán rằng có nên bật/tắt motor nước hay không **var4(t) .**Nhóm quyết định sử dụng n=3.

Sau đó ta tiếp tục chia train\_data thành 2 tập : tran và validation:

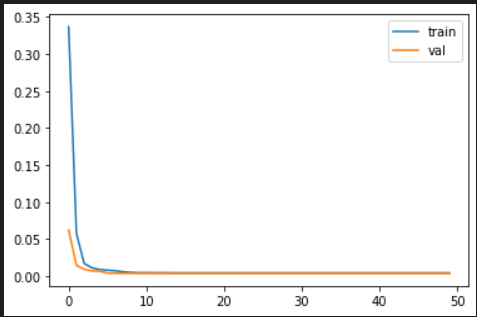


Tiếp theo nhóm tiếp tục xây dựng model LSTM:

Mô hình gồm 5 layer. Trong đó 4 layer đầu là mạng LSTM, với mỗi layer sử dụng 50 cell LSTM, và sau mỗi layer là một lớp dropout để giúp mô hình tránh khỏi hiện tượng overfitting. Layer cuối cùng là lớp kết nối đầy đủ (fully connected layer) và hàm kích hoạt được sử dụng sigmoid rất phù hợp với những bài toán phân loại nhị phân

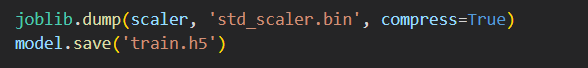


Kết quả train mô hình :



Dựa vào hình trên ta thấy loss train và validation giảm đều và có thể hội tụ trước 50 epoch

Cuối cùng ta lưu mô hình sau khi đã train để có thể tái sử dụng lại :



Tltk:

<https://machinelearningmastery.com/convert-time-series-supervised-learning-problem-python/>

<https://machinelearningmastery.com/multivariate-time-series-forecasting-lstms-keras/>