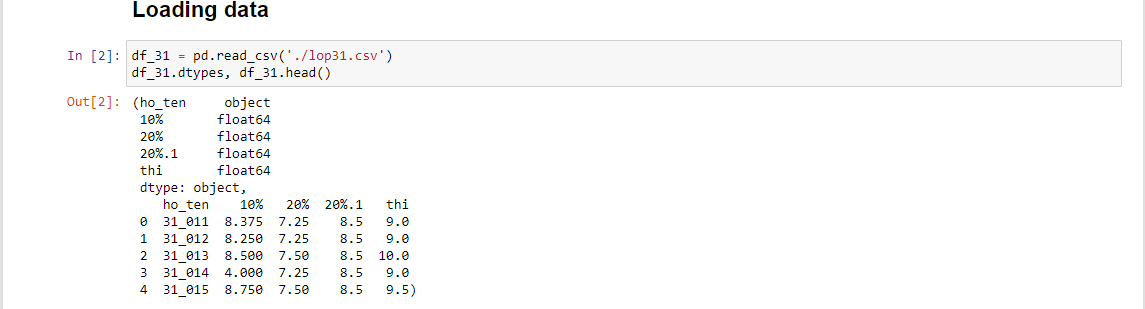
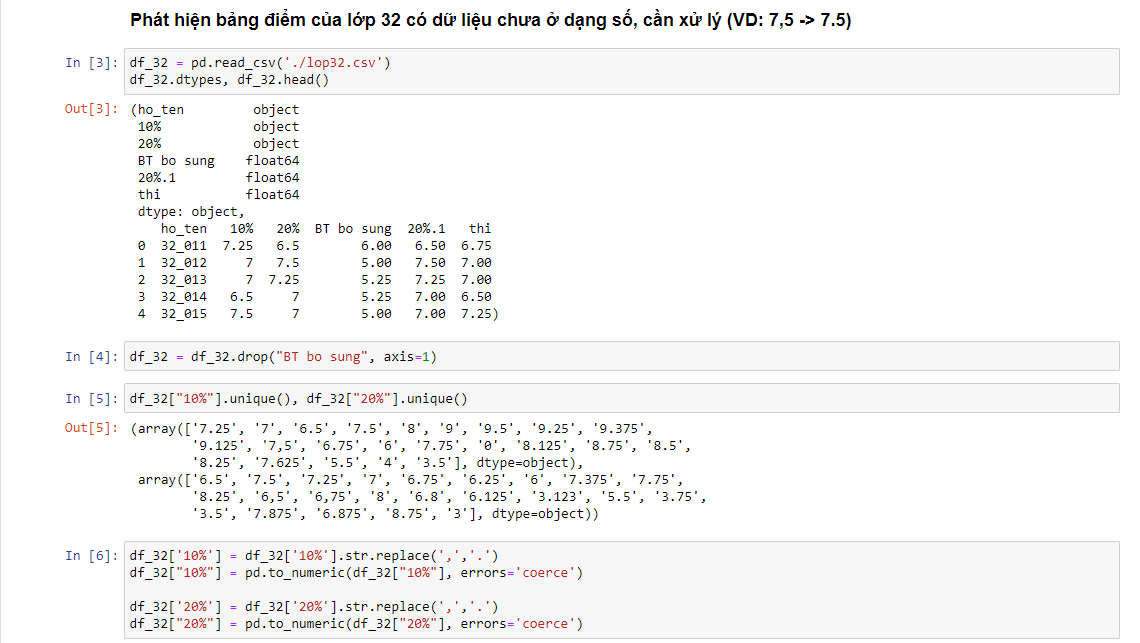
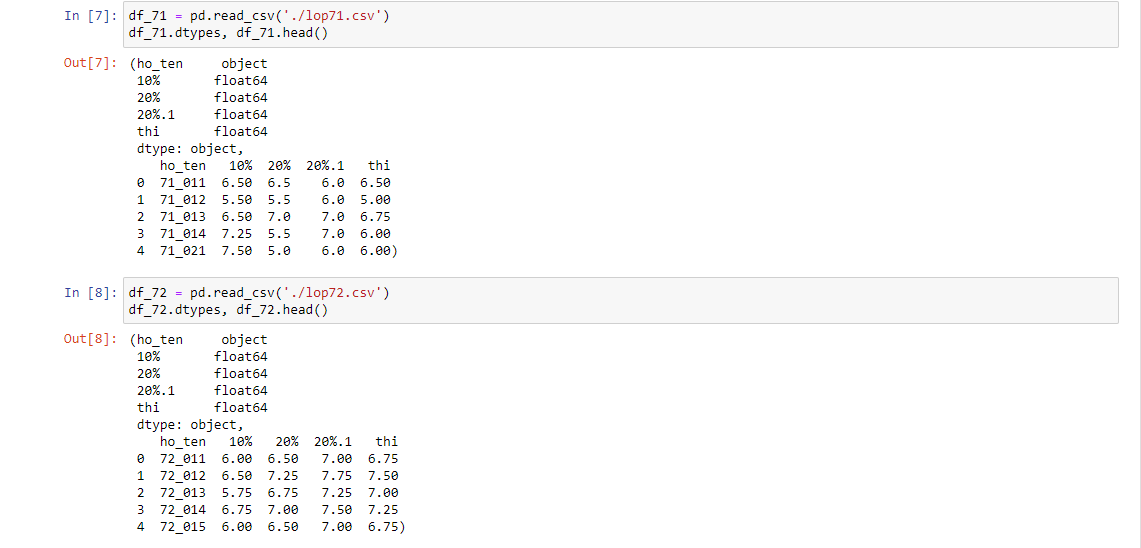
Case study 2: Xây dựng mô hình dự đoán điểm thi của sinh viên dựa trên điểm thành phần

1. Xử lí dữ liệu

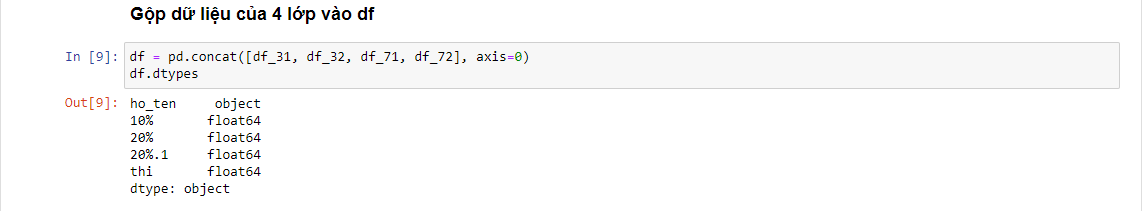
* Đọc dữ liệu từ file csv và làm sạch dữ liệu: gồm bảng điểm của 4 lớp: lop71, lop72, lop31, lop32



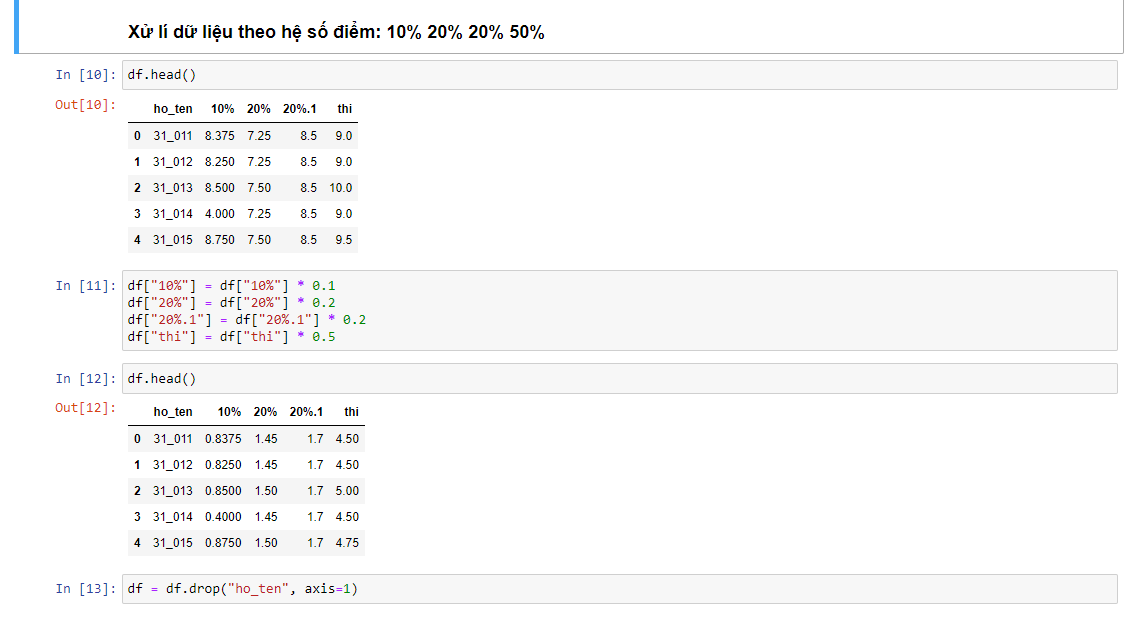




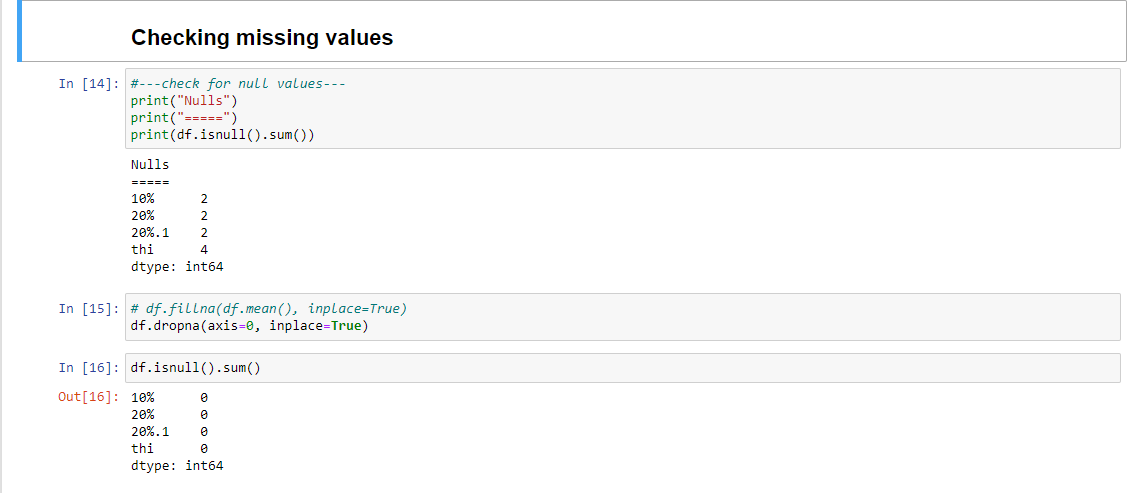
* Gộp dữ liệu của 4 lớp vào một DataFrame



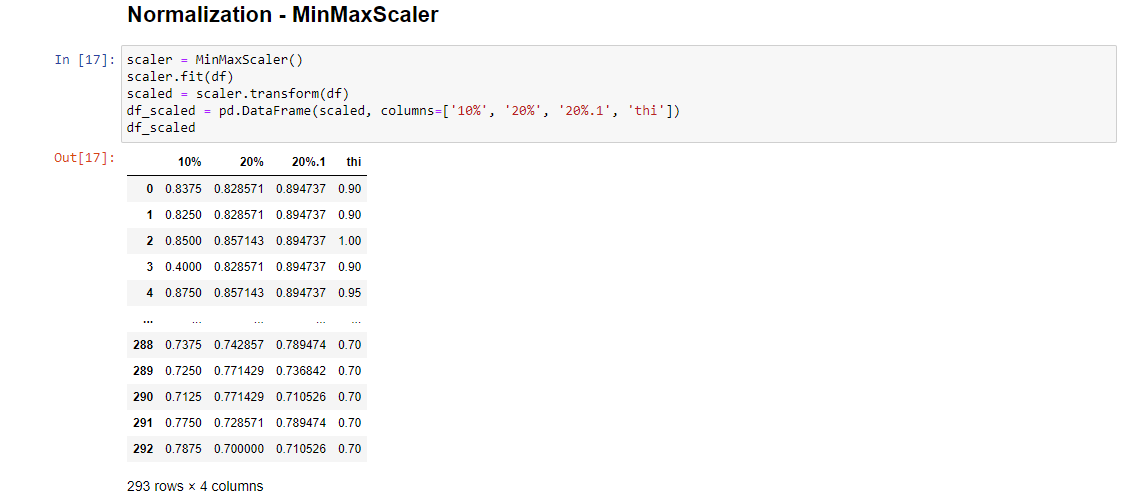
* Xử lí dữ liệu theo hệ số điểm: 10% 20% 20% 50%



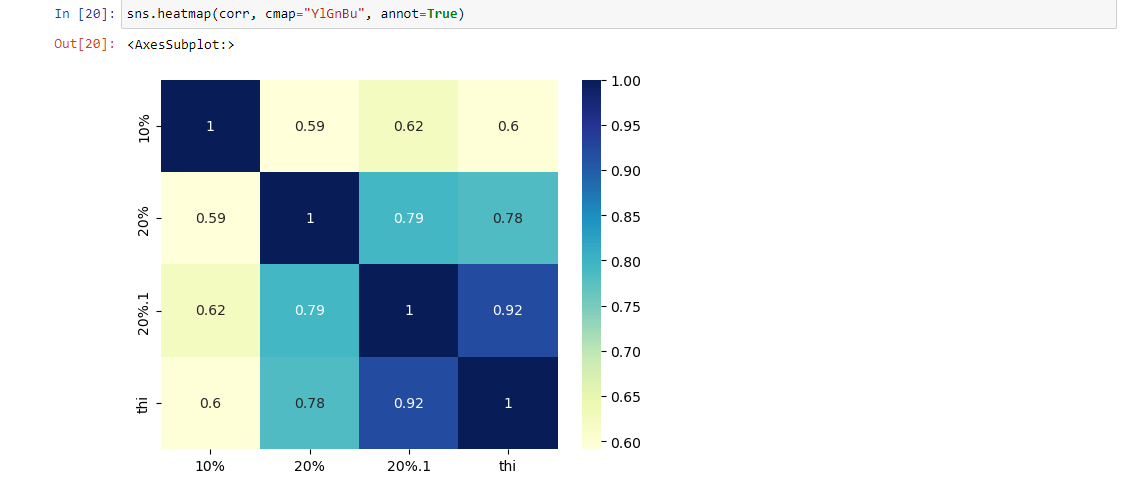
* Loại bỏ những dòng thiếu dữ liệu



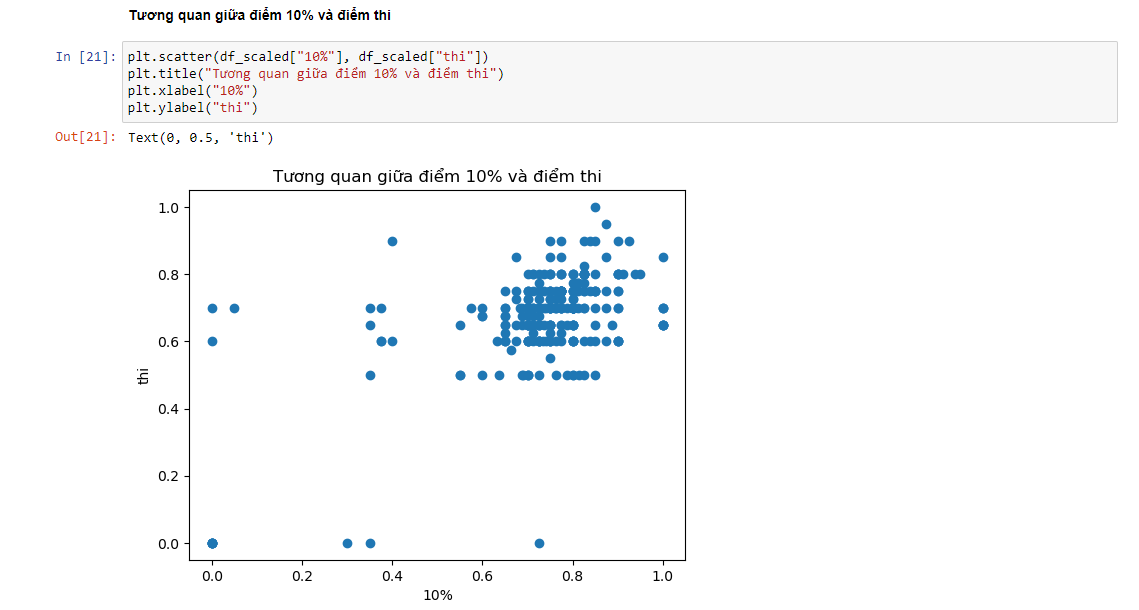
* Normalization: MinMaxScaler() -> dữ liệu sẽ nằm trong khoảng (0, 1)

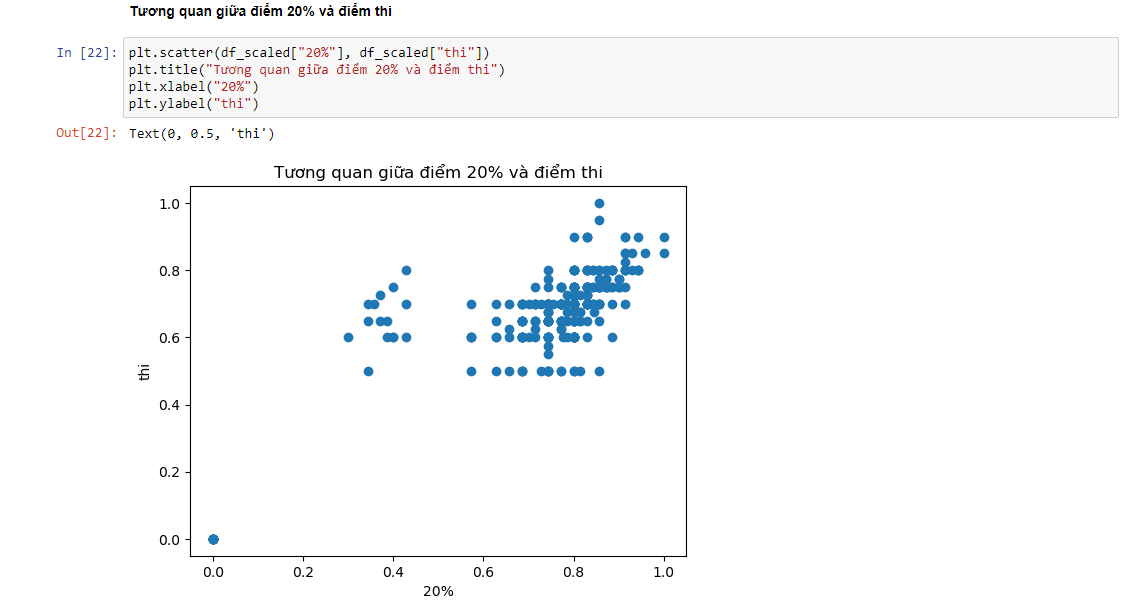


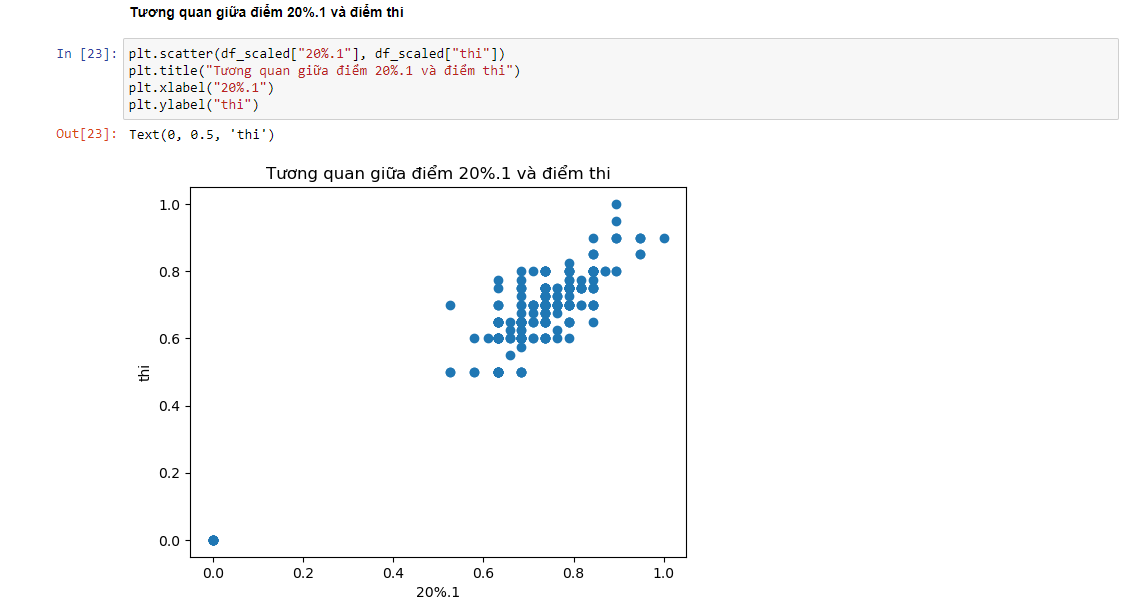
* Xem độ tương quan giữa các features



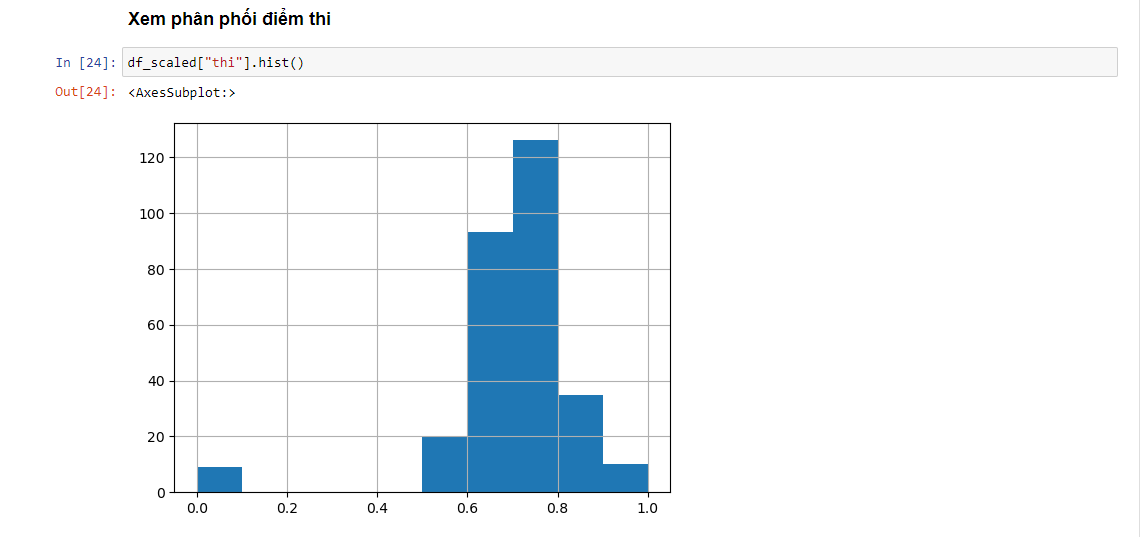
* Trực quan hóa dữ liệu





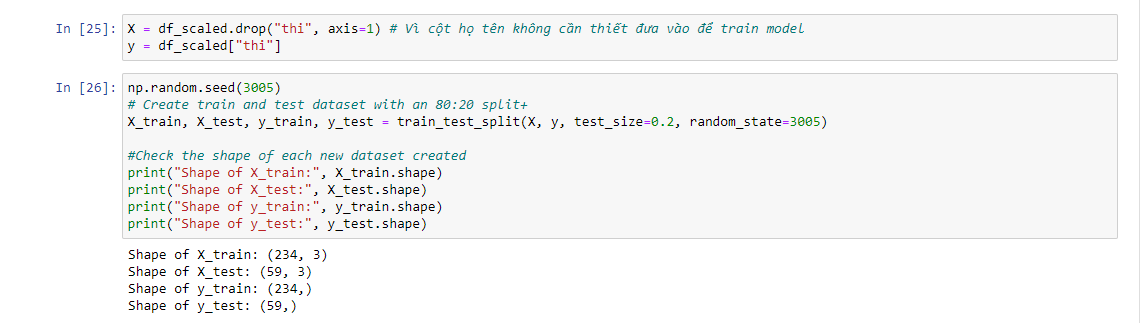


* Xem phổ điểm

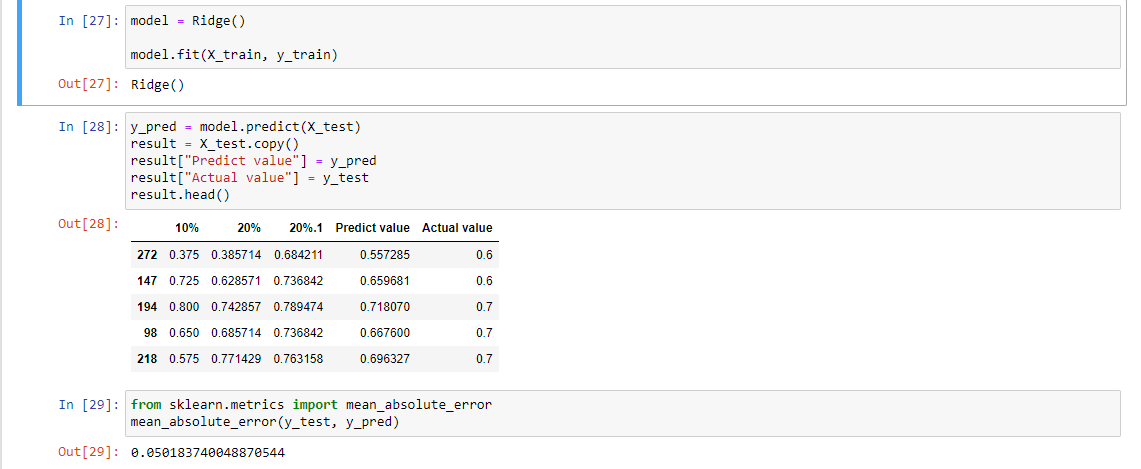


1. Mô hình machine learning

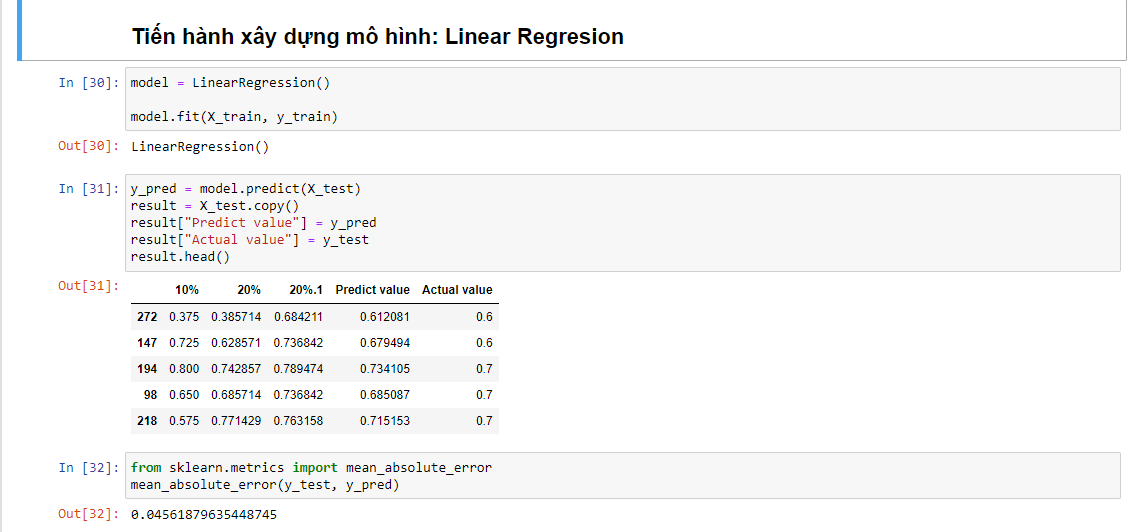
* Chia dữ liệu thành 2 tập train & test với tỉ lệ 80:20



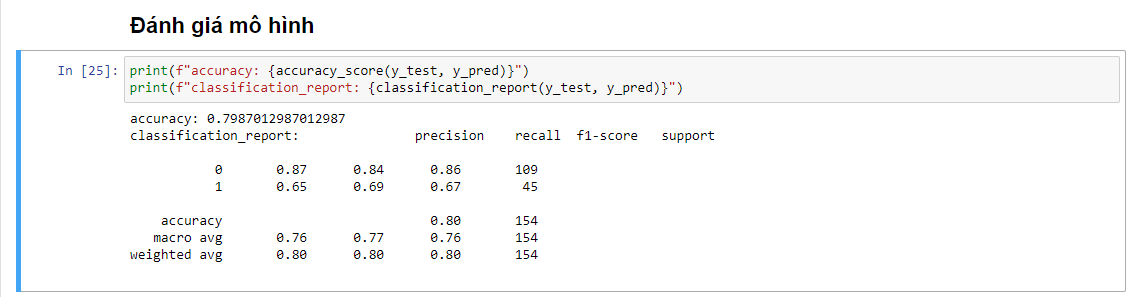
* Tiến hành xây dựng và đánh giá mô hình: Ridge Regresion



* Tiến hành xây dựng và đánh giá mô hình: Linear Regresion

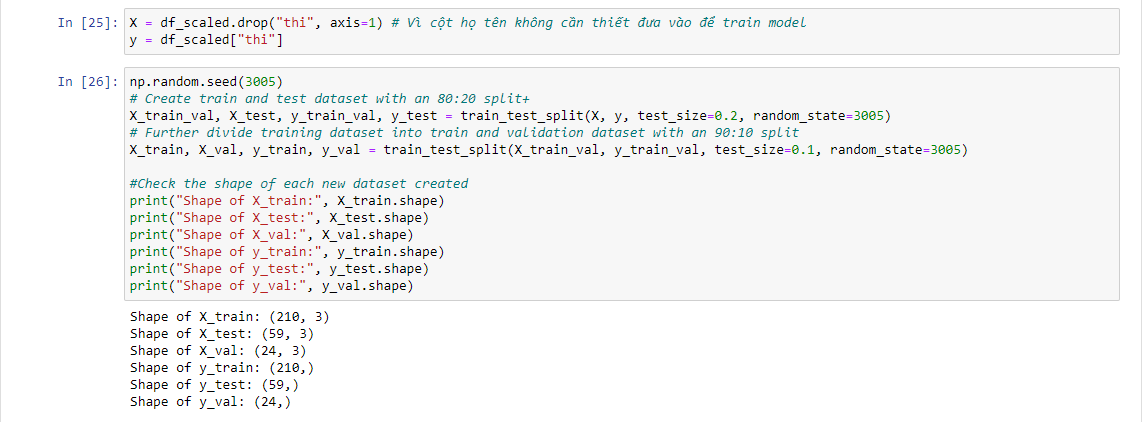


* Đánh giá mô hình:

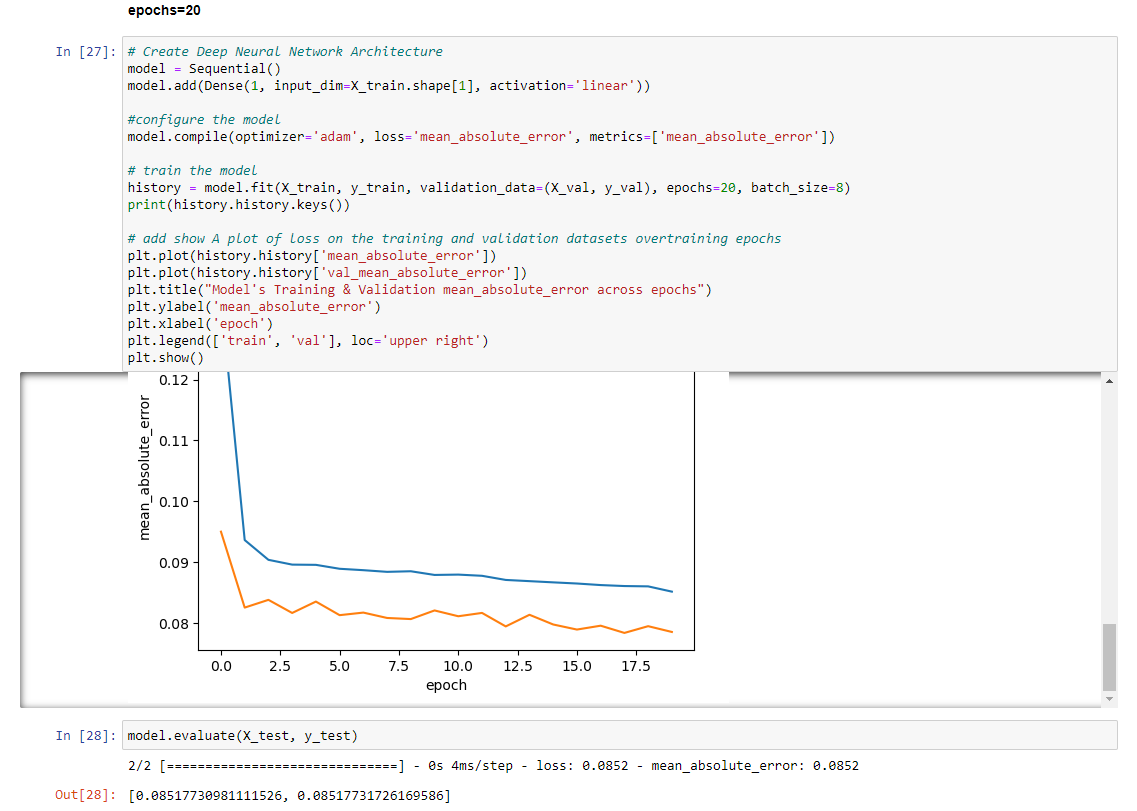


1. Mô hình deep learning

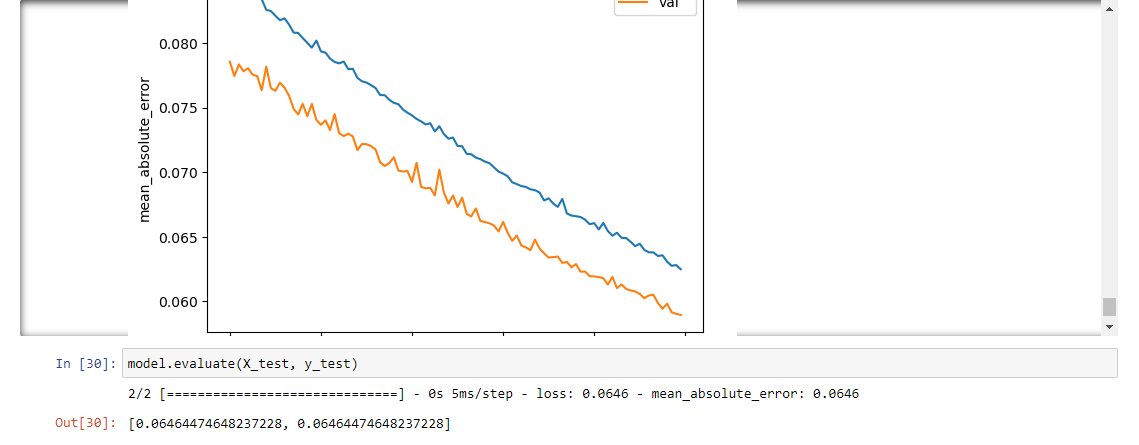
* Chia dữ liệu thành 3 tập train, val, test



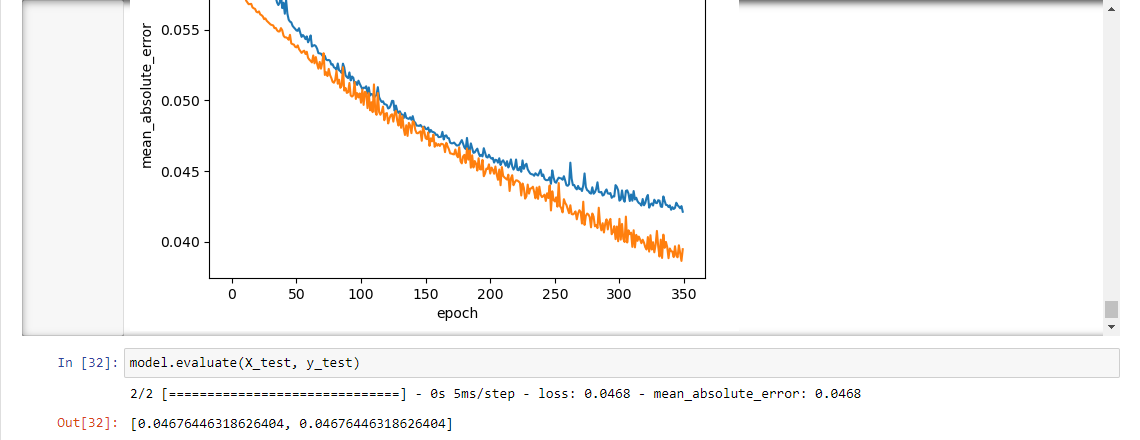
* Xây dựng mô hình Deep Learning với các tham số: *optimizer='adam', loss='mean\_absolute\_error', metrics=['mean\_absolute\_error']*
* 0 hidden layer - 1 neurons, batch\_size=8
  + epochs=20



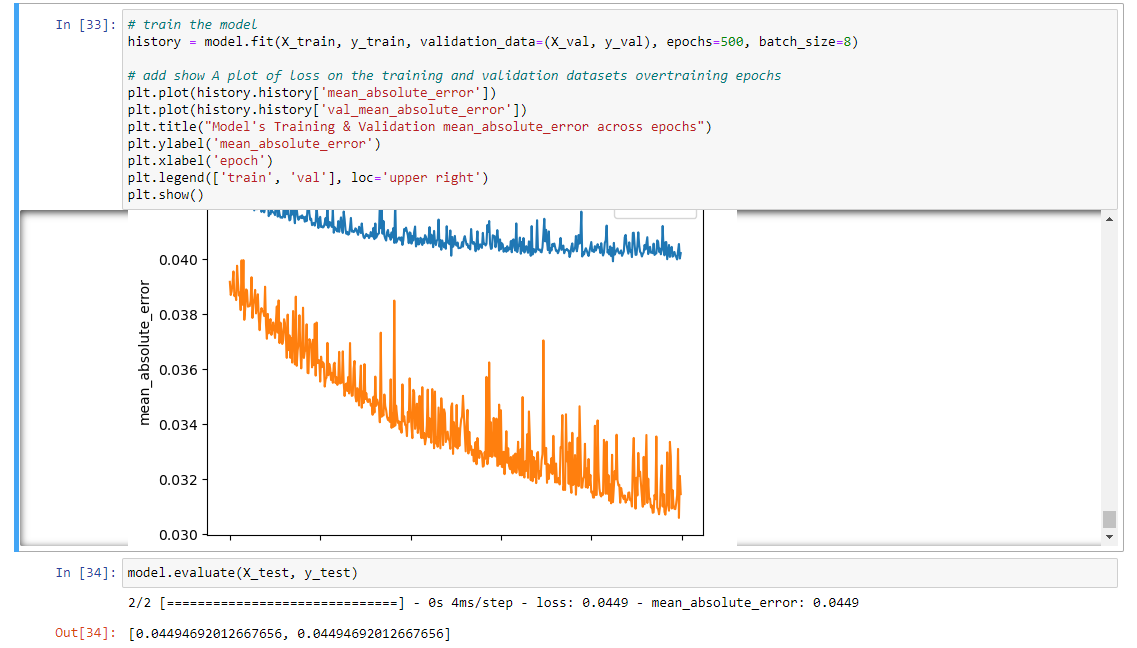
* + epochs=100



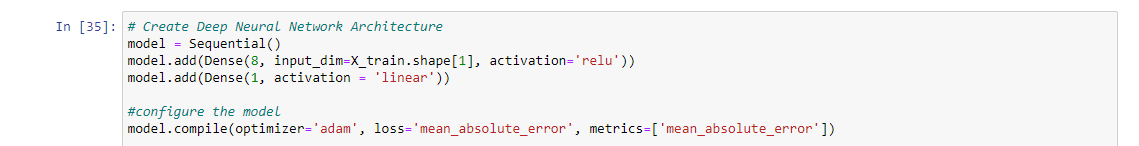
* + epochs=350

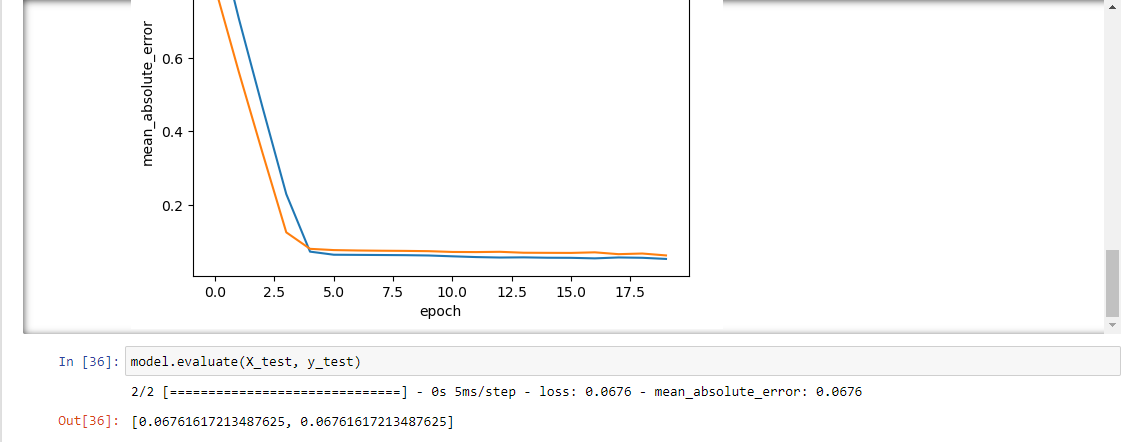


* + epochs=500

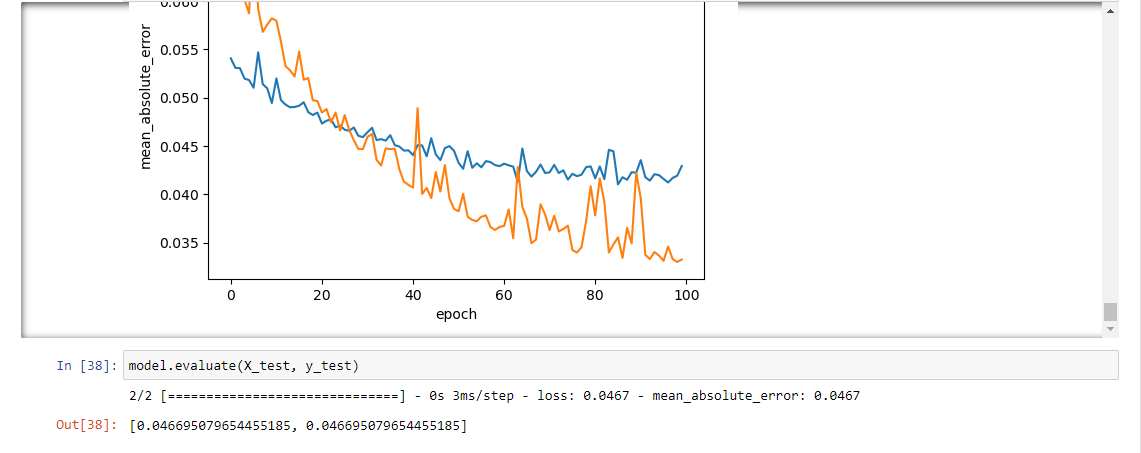


* 1 hidden layer - 8 neurons, batch\_size=8
  + epochs=20

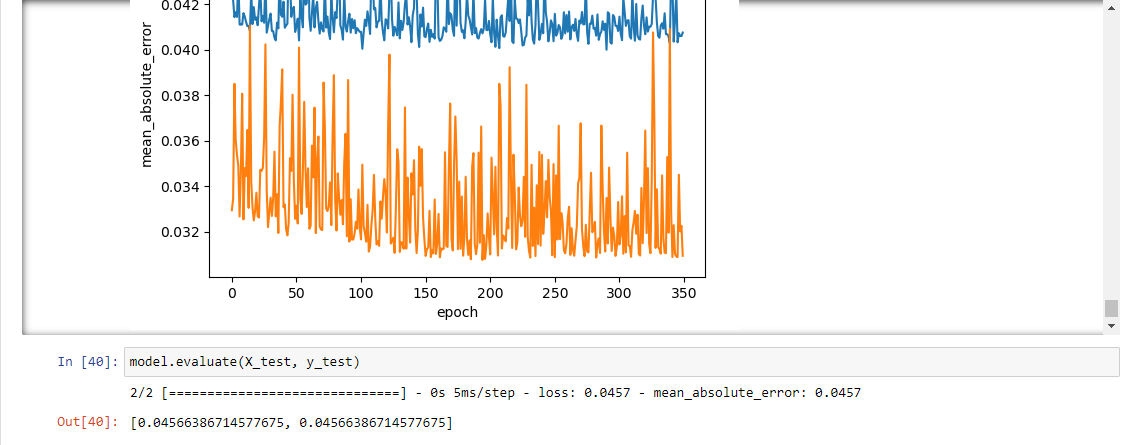




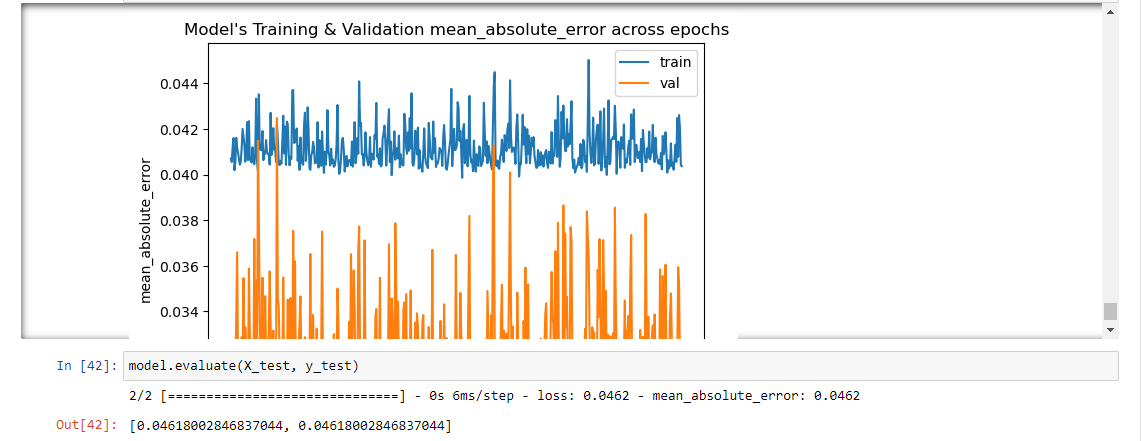
* + epochs=100



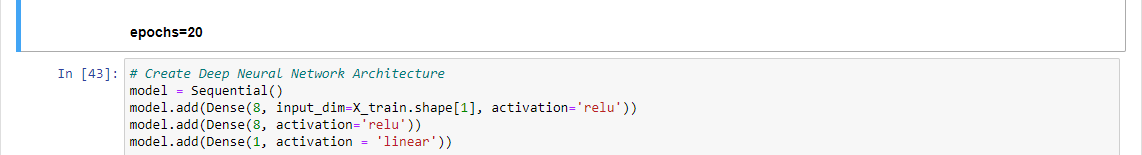
* + epochs=350

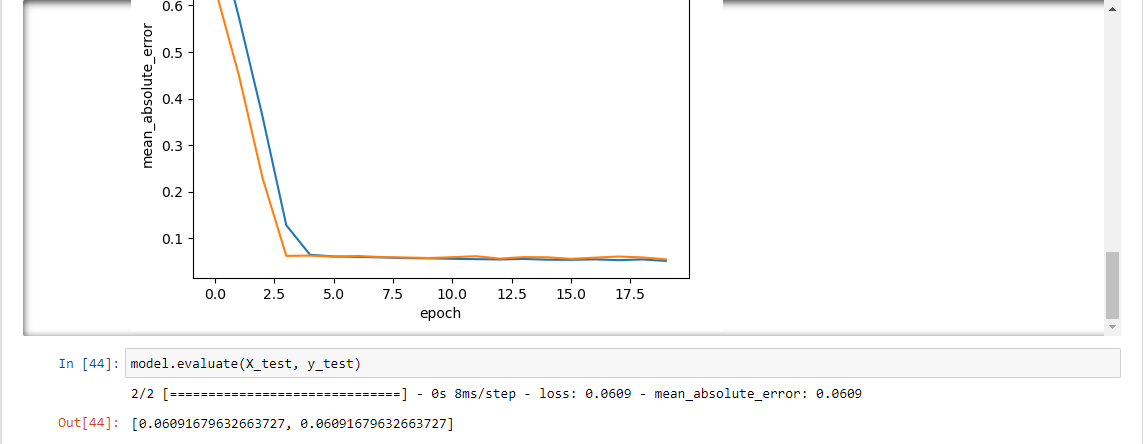


* + epochs=500

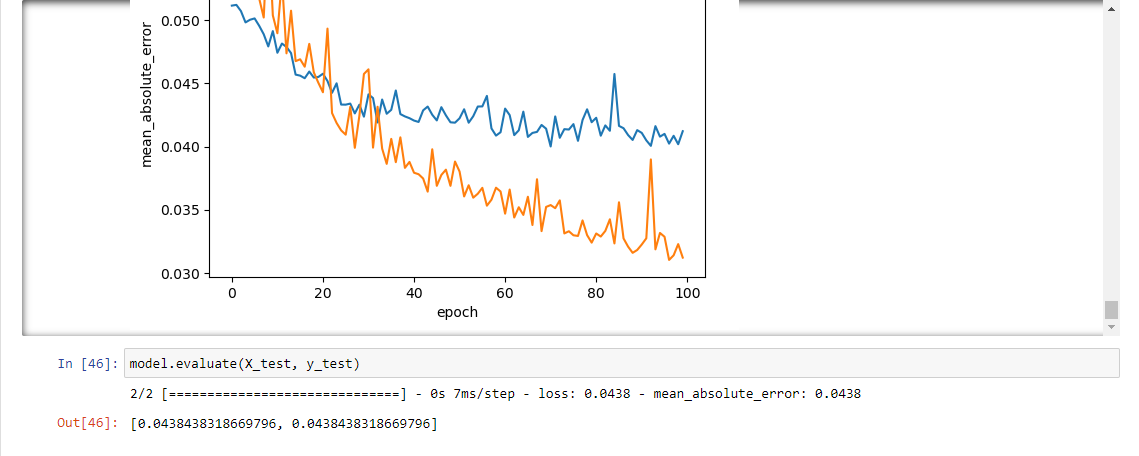


* 2 hidden layer - 8 neurons, batch\_size=8
  + epochs=20

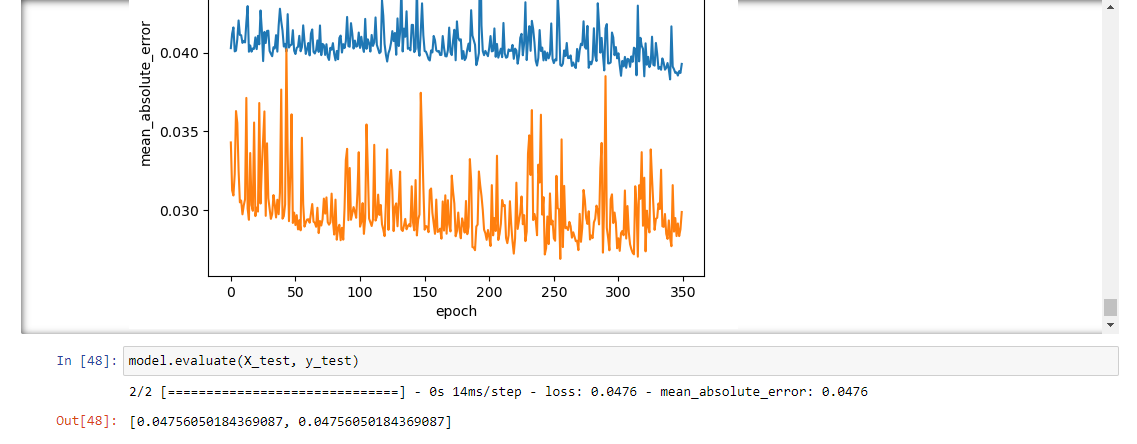




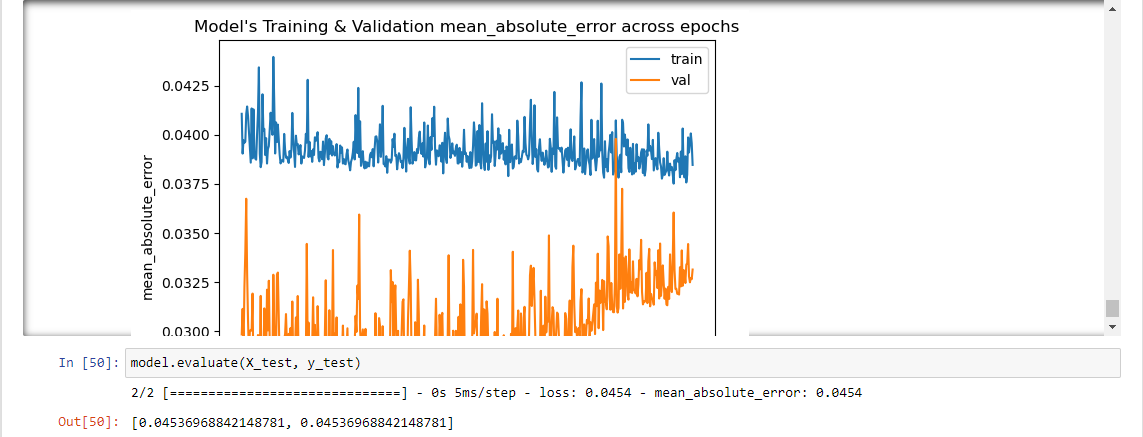
* + epochs=100



* + epochs=350



* + epochs=500



* Deepest - 2 hidden layer - 16 neurons, batch\_size=8, epochs=2000

