Cho tập dữ liệu hãy thực hiện các bước phân tích sau: Xây dựng mô hình hồi quy tuyến tính dự báo điểm thi LOGIC dựa trên các điểm toán và giới tính trong các học kì của năm cấp 3

1. Tải dữ liệu diemthidaihoc.csv và hiển thị các dòng dữ liệu
2. Tiến hành EDA dữ liệu, ví dụ: Hãy thống kê các độ đo cơ bản của dữ liệu và nhận xét
3. Chia tập dữ liệu thành tập train và test với tỉ lệ 80:20 và random state là 16
4. Tiến hành transform theo hàm x\*x dữ liệu input đầu vào trên tập train
5. Tiến hành EDA dữ liệu các điểm toán sau khi transform. Ví dụ: Hãy vẽ biểu đồ phân phối các điểm toán và cho nhận xét
6. Tiến hành tiếp tục chuẩn hóa minmaxscale dữ liệu đầu vào trên tập train
7. Tiến hành EDA dữ liệu các điểm toán sau khi minmaxscale. Ví dụ: Hãy vẽ biểu đồ Boxplot của điểm toán các học kì và cho nhận xét
8. Xây dựng mô hình hồi quy tuyến tính dự báo điểm LOGIC dựa trên các yếu tố điểm toán và giới tính
9. Hãy đánh giá mô hình trên tập train và test. Ví dụ: so sánh kết quả MAE trên tập train và test đưa ra nhận xét
10. Ứng dụng dự báo, với điểm toán các học kì 1,2 các năm lớp 10,11,12 và giới tính lần lượt là: 5.0, 6.5, 7.0, 7.0, 6.5, 6.0 và giới tính là nam thì điểm thi LOGIC là bao nhiêu.