1. Kiểm tra lại danh mục hình ảnh

2. Danh mục bảng biểu chưa có

3. Phần thu thập và mô tả dữ liệu em vẽ một số hình mô tả dữ liệu ví dụ dưới dạng line chart, hoặc đưa ra các bảng thống kê về dữ liệu.

4. Trong mỗi bước tiền xử lí dữ liệu, em nêu rõ sẽ áp dụng kĩ thuật này cho những cột nào.

5. Phần trực quan hóa dữ liệu thì nên để tiêu đề là khai phá dữ liệu.

Sau bước khai phá dữ liệu này, em đặt câu hỏi bài toán là gì? Em phải thêm cái dòng kết luận là Muốn xây dựng một bài toán dự đoán giá nhà dựa và các mô hình học máy. Thì như vậy mới có đến mục 6. là Sử dụng mô hình học máy.

6. Phần 5 và 6 này chưa hợp lí, em nên gộp nó lại thành 1 phần là:

5. Thực nghiệm dự đoán giá nhà sử dụng học máy (ML)

5.1 Lựa chọn feature (Đây là phần mà em áp dụng SelectKBest,...)

5.2 Các mô hình học máy (Đây là nơi em viết sơ qua về lí thuyết các mô hình này)

a) Linear Regression

b) Polynomial Linear Regression.

c) Random Forest

5.3 Các độ đo hiệu năng (Đây là nơi em viết sơ qua lí thuyết về các hàm đo hiệu năng MSE, MAE, RMSE, R2 score)

5.3 Kết quả và đánh giá

Em tổng hợp kết quả cả 3 mô hình trên vào 1 bảng duy nhất. Sau đó nhận xét và đưa đánh giá. Ngoài ra, vẽ các hình kết quả lên, ví dụ giá trị thực tế và giá trị dự đoán.

==> Sau đó kết luận chung là mô hình nào phù hợp nhất cho bài toán này.