

BÀI TẬP SỐ 1

Lớp chia nhóm, mỗi nhóm 3 sinh viên, mỗi nhóm chọn một đề tài để thực hiện phương pháp Hồi quy tuyến tính. Yêu cầu thực hiện của mỗi đề tài gồm:

1. Mô tả đề tài: Tên đề tài, mô tả tóm tắt công việc thực hiện của đề tài.
2. Mô tả tập dữ liệu của bài toán: Dữ liệu gồm những chiều thông tin gì (mỗi vector dữ liệu có những thông tin gì), có bao nhiêu vector dữ liệu (ít nhất là 100 vector dữ liệu). Ví dụ dự đoán về bệnh tim của một người, thì mỗi người có 1 vector dữ liệu gồm các thông tin: tuổi, giới tính, nhịp tim, huyết áp, Thông tin cần dự đoán của bài toán là gì (mô tả dữ liệu đầu ra tương ứng với mỗi vector dữ liệu).
3. Chia tập dữ liệu thành 2 phần: 70% dùng để huấn luyện mô hình, 30% dùng để kiểm tra sự phù hợp của mô hình.
4. Mô tả ma trận dữ liệu huấn luyện (X), vector đầu ra (Y)
5. Dùng phương pháp Hồi quy tuyến tính để xây dựng mô hình cho bài toán với tập dữ liệu huấn luyện.
6. Dùng tập dữ liệu kiểm tra để so sánh kết quả dự đoán của mô hình với kết quả thực tế, để đánh giá sự phù hợp của mô hình.

Lưu ý:

- Nộp bài: Các nhóm upload tập dữ liệu của đề tài (gồm: file word trình bày bài tập, các file dữ liệu khác để thực hiện đề tài)
- Hạn nộp: trước 23h00 thứ 5 ngày 02/12/2021