N chiều dữ liệu thì có n x (Xn)

Tính đạo hàm từng biến 1 để tìm ra tham số tối ưu

Có 2 loại normalization

Min Max sxaling ?

Mean nỏmalization

- dữ liwuwj k tuyến tính : lecture 8

Thầy nói? Nhiệm vụ tìm số chiều sao cho tuyến tính

Giả thuyết: giải dụ 1 chiều không gian A là tuyến tính rồi tính tổng mất mát

=> tổng mất mát nhỏ hơn kgian 2 chiều => tăng chiều có ích

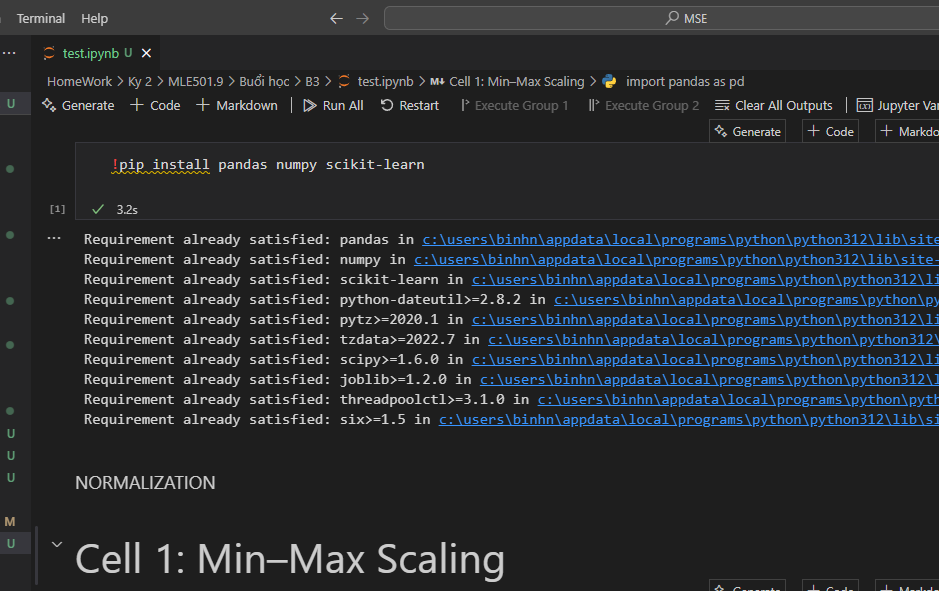
=> khi nào tăng và giảm chiều

Mã hóa nghịch đảo ? RSSA 256

À ví dụ Furier rtanform thì cí inverrt FT

Hoặc nén và giải nén và ngc lại nó là bản chất của chiếu 2 chiều (ít chiều chiếu lên nhiều chiều và ngc lại)

Kaggle hay tải dử liệu mẫu



Trên là xử lý dữ liệu trc khi đưa vào mô hình

Ném normalition vào liner regersionn xem có giảm loss function k

LECTURE 6: phân lớp

Cách chia dữ liệu sao cho hợp lý

Tránh oveerrfitiing bằng 2 bài toán sau:

Thường à tăng thêm data sẽ ngon hơn

THuật toán KNN - láng giềng -> dùng K a láng giềng gần nhất để dự đoán dữ liệu của label đó,

=> làm khoa học phải tính độ phức tạp của thuật toán => thử nghiệm với iris

Tổng kết:

Phân loại classsication => overrfitiing => KNN=> đọc dữ liệu