

Họ và tên: Nguyễn Minh Phú

MSSV: 19520218

Môn học: CS116.M11.KHTN

Phân biệt bài toán Hồi quy (Linear Regression) và bài toán Phân lớp (Classification)

Điểm chung là cả 2 bài toán đều thuộc dạng Supervised Learning. Điểm riêng duy nhất là loại giá trị trả về (output)

Hồi quy Tuyến tính (Linear Regression)

Bài toán Hồi quy sẽ trả về giá trị (output) **liên tục** (continuous), **1 con số** dự đoán, và những điểm giá trị trả về này có **tính thứ tự**, có thể **so sánh** bằng các phép tính $<$, $>$, \leq , \geq . Nhiệm vụ của bài toán là tìm 1 đường cong, thẳng, hoặc các mặt phẳng hay siêu mặt phẳng để phù hợp với phân phối dữ liệu để xấp xỉ đầu ra.

Ví dụ bài toán dự đoán giá nhà dựa trên những đặc điểm của căn nhà đó: diện tích đất, đất có thổ cư, diện tích sử dụng,.... Thì khi đưa vào 1 mẫu dữ liệu tương ứng thì ta có thể đoán được giá gần đúng thực tế của ngôi nhà ở ngoài đời thật.

Phân lớp (Classification)

Bài toán Phân lớp sẽ trả về 1 giá trị (output) là **nhãn của dữ liệu** (classes), 1 con số tương ứng cho nhãn của dữ liệu. Tuy nhiên, những giá trị được trả về là **rời rạc** (discrete) và không có tính thứ tự, chỉ có thể so sánh bằng phép tính $=$, $<>$. Nhiệm vụ của bài toán là tìm những đường ranh giới của các lớp dữ liệu để chia dữ liệu thành các lớp khác nhau.

Ví dụ bài toán phân lớp giới tính thông qua ảnh, đầu vào là 1 tấm ảnh, thì ta sẽ biết được giới tính của người đó.

Bài toán Dự đoán tuổi từ gương mặt

Đây là bài toán hồi quy. Lý do là đầu ra của bài toán là tuổi, có thể chính xác hoặc không, nhưng đại lượng tuổi có thể xem là liên tục, có tính thứ tự, và so sánh được thông qua các dấu bé, lớn...