

Neural Network

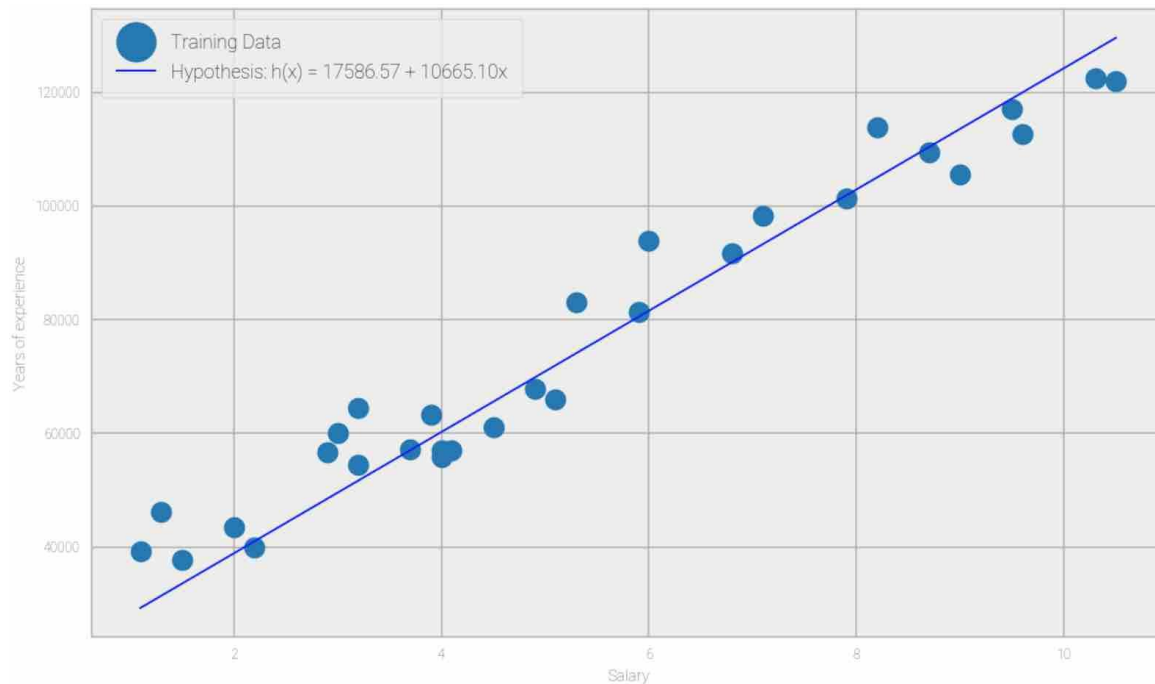
FIC - Ta Dang Khoa

Nội dung

- Nhắc lại các mô hình đã học
- Neural network
- Backpropagation

Nhắc lại về các mô hình đã học

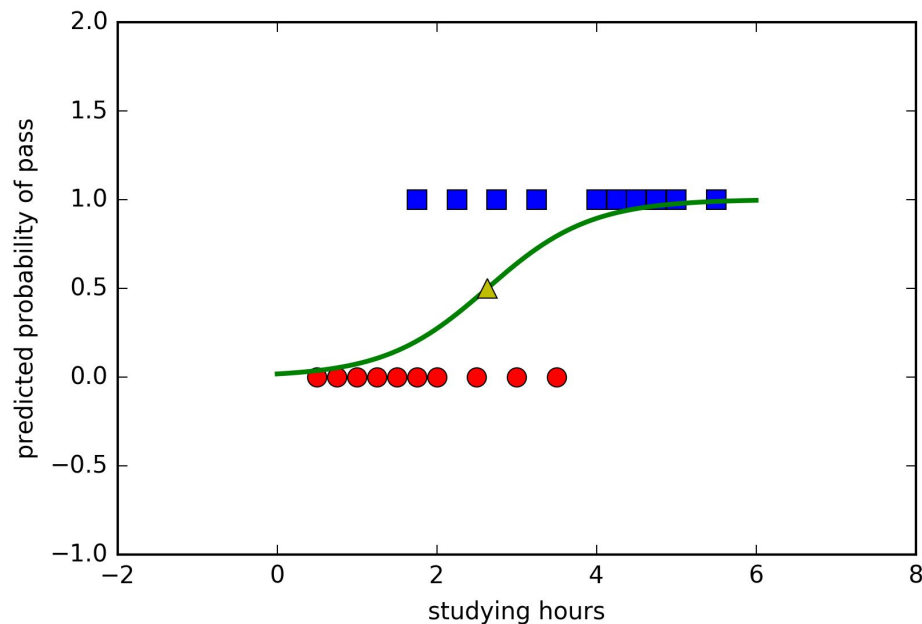
Linear regression



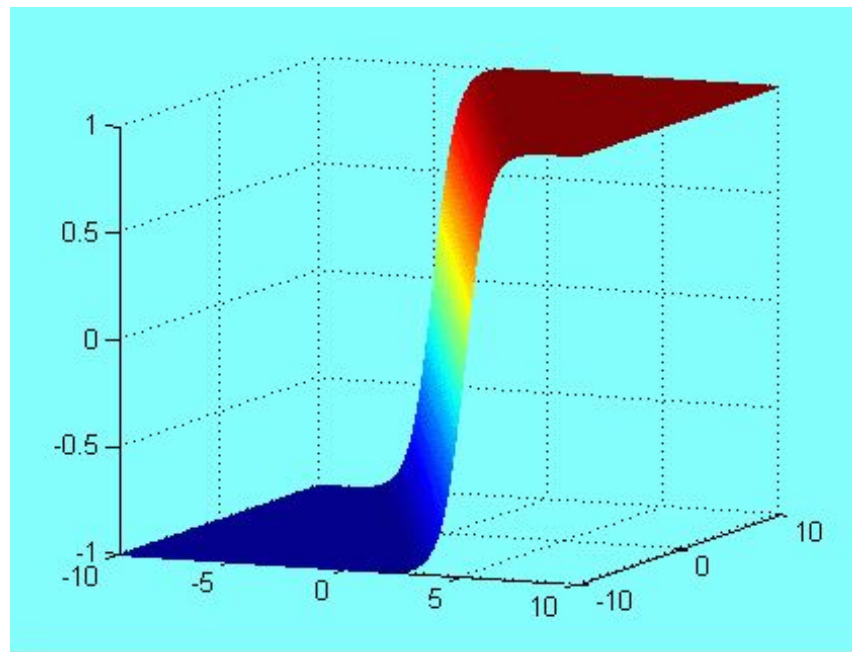
Bài toán dự đoán giá nhà dựa trên diện tích

Nhắc lại về các mô hình đã học

Binary classification



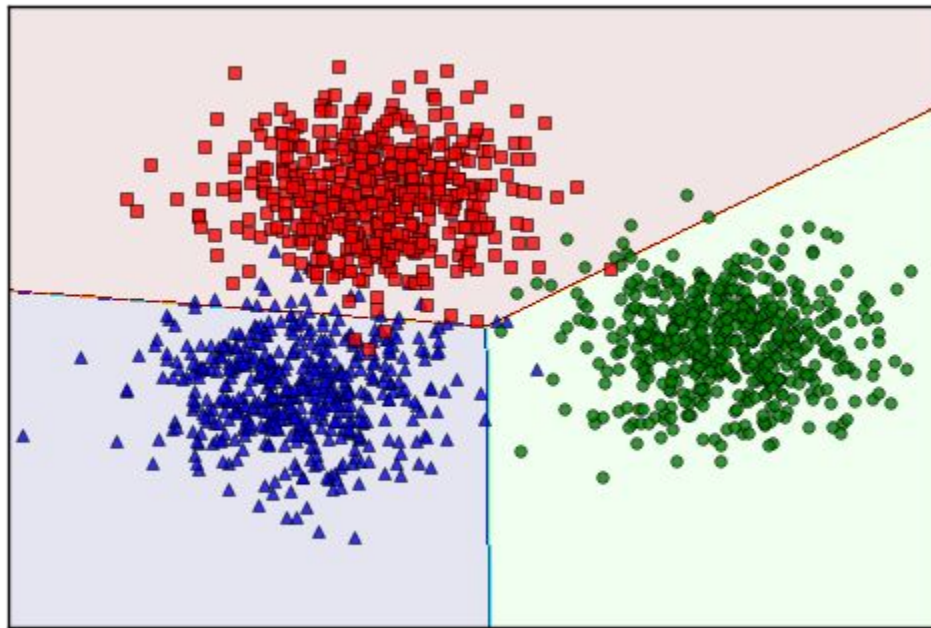
Sigmoid với dữ liệu 1 chiều



Sigmoid với dữ liệu 2 chiều

Nhắc lại về các mô hình đã học

Multiclass classification



Ranh giới giữa các class tìm được bằng Softmax Regression. (Nguồn machinelearningcoban.com)

Nhắc lại về các mô hình đã học

Kết luận

- Dữ liệu đơn giản
- Hàm dự đoán là tổ hợp tuyến tính của các feature
- Dữ liệu phân tách tuyến tính rõ ràng

=> những bài toán có data phi tuyến tính thì model khó (không) học được

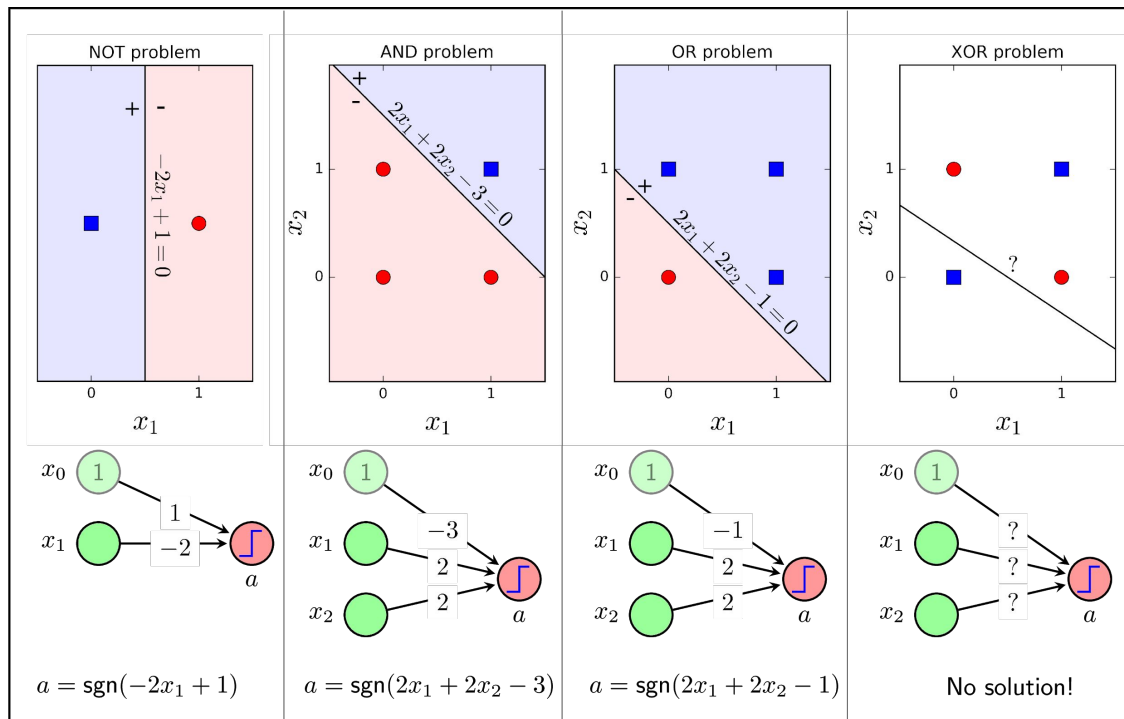
Neural network

Định nghĩa

- Dạng cơ bản nhất của neural network (**linear model**) là tổ hợp tuyến tính của input
- Mạng deep neural network kết hợp nhiều các **`linear model`** với các **`hàm phi tuyến`** và **`cấu trúc phân lớp`**

Neural network

Ví dụ

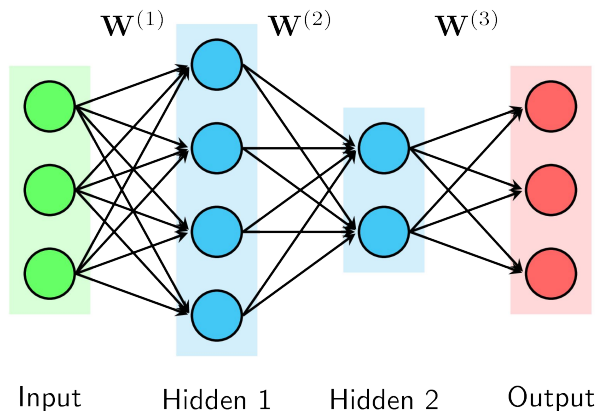


Ví dụ sử dụng 1 tổ hợp tuyến tính và một hàm phi tuyến (sigmoid) (Nguồn: machine learningcoban.com)

Neural network

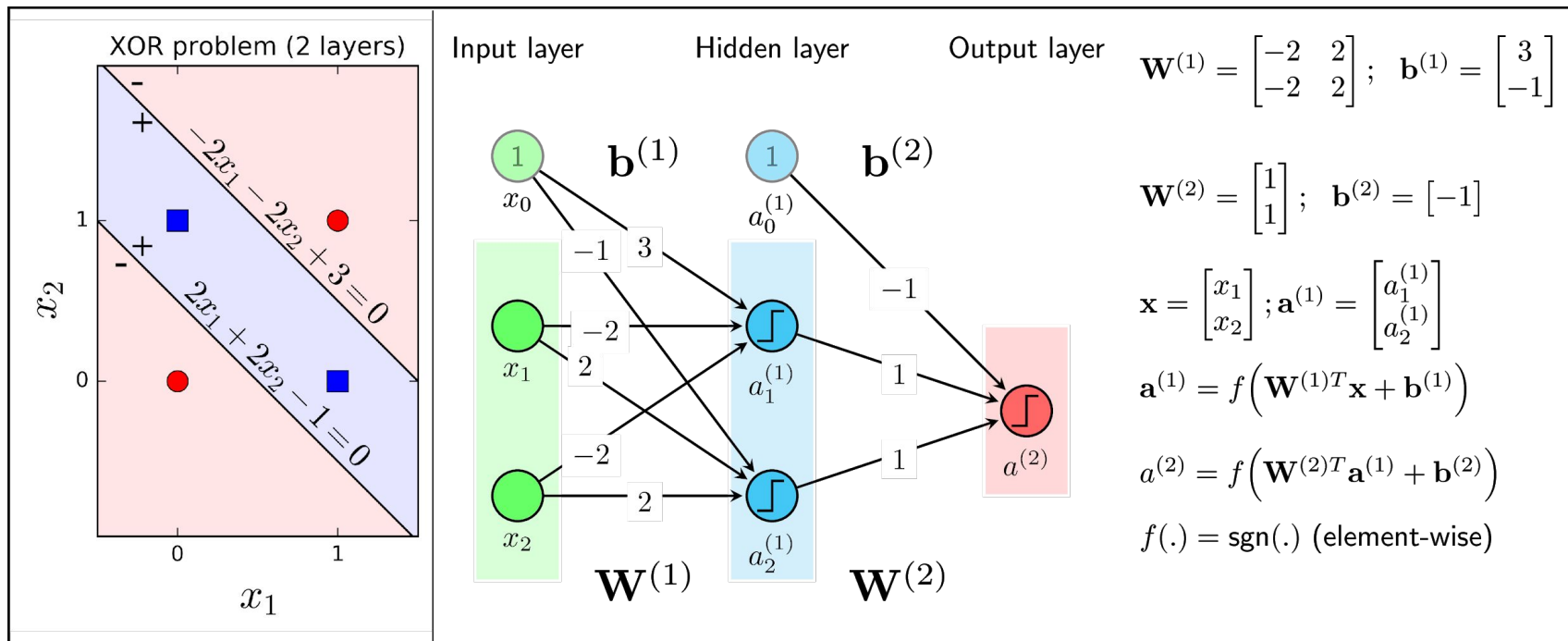
Tại sao tốt?

- Hàm phi tuyến bề cong không gian của dữ liệu => phân tách tuyến tính
- Sử dụng nhiều hidden layer giúp:
 - Cách hiểu 1: Mỗi unit thì tương đương với một feature mới. Sử dụng nhiều hidden layer giống như việc cô đọng ý nghĩa các feature (VD: cộng nhiều thành nhân, nhân nhiều thành mũ ...)
 - Cách hiểu 2: Tăng số hidden layer => tăng độ phức tạp của hàm dự đoán => Tăng khả năng fit được với dữ liệu



Neural network

Ví dụ

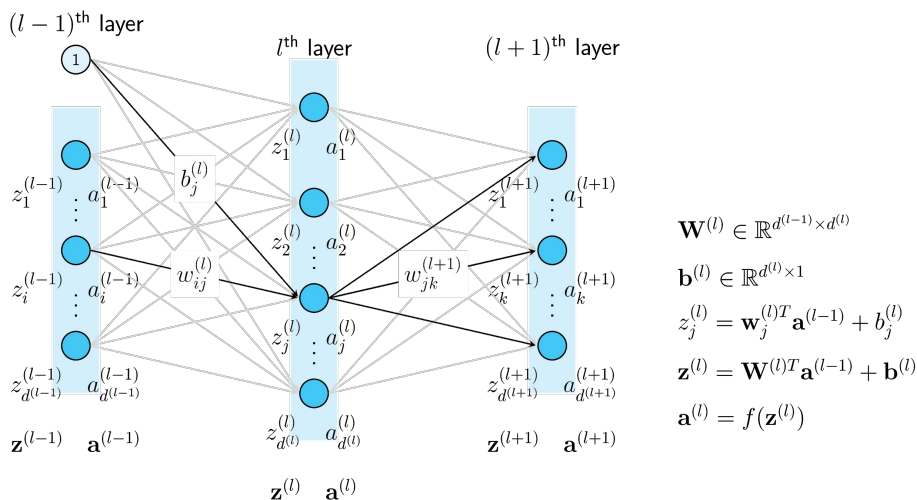


Ví dụ sử dụng deep NN (Nguồn: machinelearningcoban.com)

Backpropagation

Giống như các mô hình trước đã học. Chúng ta cũng tối ưu hóa hàm loss bằng việc sử dụng gradient descent.

=> Thực hiện tính đạo hàm theo tham số như nào???



Backpropagation

Partial derivative

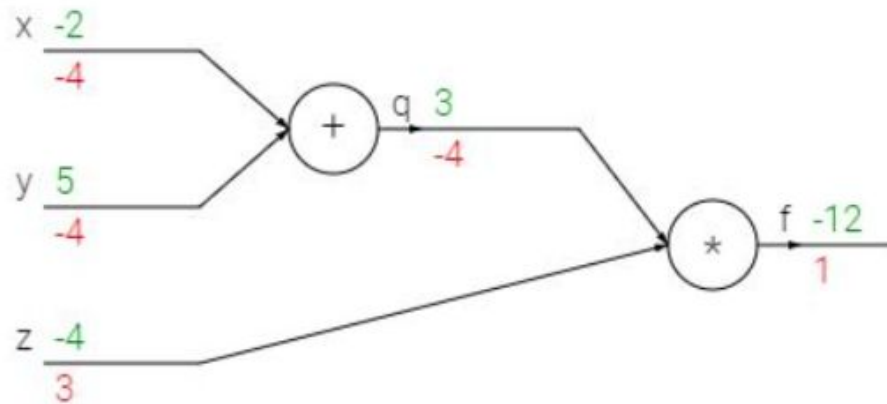
$$f(x, y, z) = (x + y)z$$

$$q = x + y \rightarrow f = qz$$

$$\frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1 \quad \frac{\partial f}{\partial q} = z = -4$$

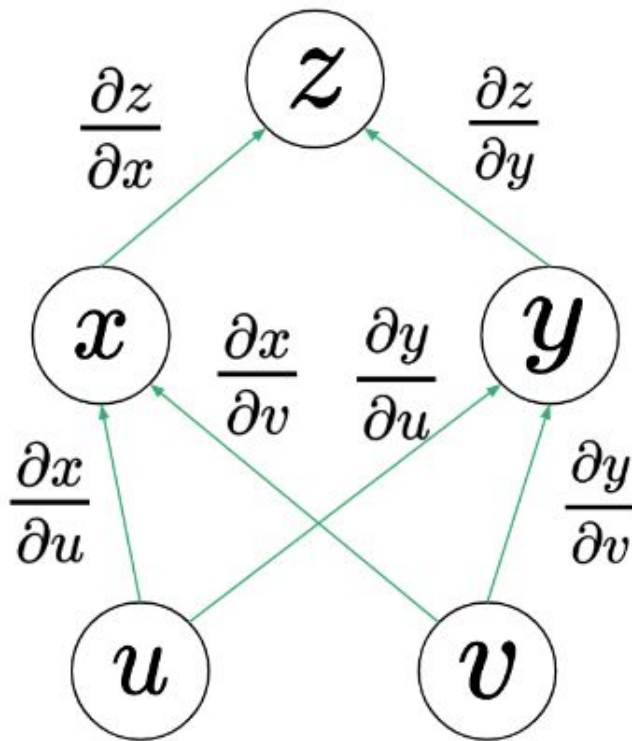
$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial q} \frac{\partial q}{\partial x} = -4$$

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial q} \frac{\partial q}{\partial y} = -4 \quad \frac{\partial f}{\partial z} = q = 3$$



Backpropagation

Multivariable chain rule



$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial u}$$

$$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial v}$$

Câu hỏi

- Nếu không áp dụng hàm phi tuyến tính thì điều gì sẽ xảy ra?
- Nếu dùng nhiều hoặc ít các hidden layer thì điều gì sẽ xảy ra?
- Tự thay công thức và đạo hàm ngược ví dụ gồm có 1 hidden layer.