

Midterm exam

Thời gian làm bài: 60 phút

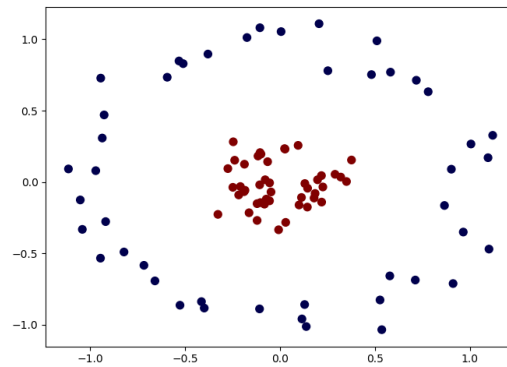
1. (3đ) Tính đạo hàm các hàm số sau (y, z là hằng số):

(a) (1đ) $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

(b) (1đ) $f(x) = \begin{cases} \alpha(e^x - 1) & \text{if } x < 0 \\ x & \text{if } x \geq 0 \end{cases}$

(c) (1đ) $f(x) = \frac{e^x}{e^x + e^y + e^z}$

2. (5đ) Cho dữ liệu như hình vẽ bên dưới:



Hình 1: Điểm đỏ tương ứng với dữ liệu label 0, điểm xanh tương ứng với dữ liệu label 1

- (a) (2đ) Thiết kế model cùng biểu diễn dưới dạng ma trận cho bài toán trên.
- (b) (1đ) Định nghĩa loss function.
- (c) (2đ) Tính đạo hàm loss function với từng biến và biểu diễn dưới dạng ma trận.

3. (3đ)

- (a) (1đ) Overfitting là gì? Cách phát hiện và giải pháp?
- (b) (2đ) Mô tả thuật toán dropout, tại sao lại hoạt động?

4. (4đ)

- (a) (1đ) Tại sao hàm activation phải non-linear?
- (b) (3đ) Tính output 1 Neural Network đơn giản (2 nodes input layer (x_1, x_2), 2 nodes hidden layer, 1 node output layer) với hàm activation $f(x) = 2x - 1$ cho tất cả các node, tất cả các hệ số $w = 1$ và $b = -1$.

5. (5đ)

- (a) (2đ) Thiết kế 1 model CNN đơn giản cho bài toán phân loại dữ liệu cifar10 (ảnh màu 32×32 , phân loại 10 lớp).
- (b) (2đ) Tính số lượng parameter ở mỗi layer ở mạng trên?
- (c) (1đ) Bài toán có input là 1 ảnh màu (32×32), output có 4 lớp (chó, mèo, cỏ, mây). Giả sử mỗi ảnh có thể thuộc nhiều hơn 1 lớp (số lượng lớp mỗi ảnh không cố định). Thiết kế model, loss phù hợp cho bài toán trên.