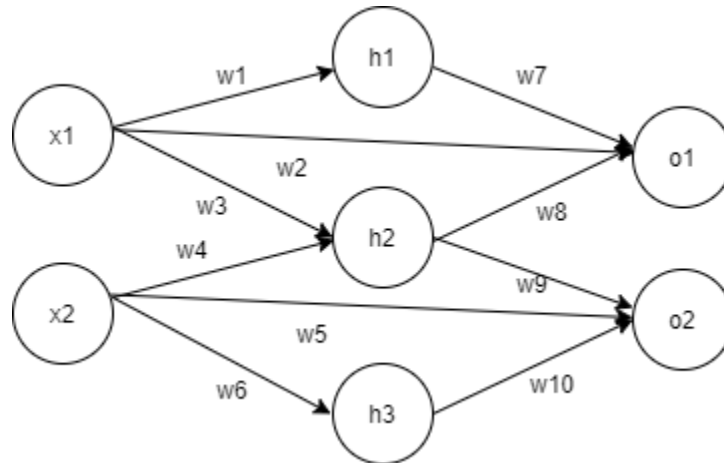


Bài 1. Câu hỏi lý thuyết (1 điểm)

1. Trình bày cơ sở lý thuyết, mục đích của tinh chỉnh L2 (L2 regularization) (0,5 điểm).
2. Giải thích sự khác nhau giữa các bộ tối ưu SGD, Adam, Adagrad, RMSProp (0,5 điểm)

Bài 2. (3 điểm) Mạng ANN

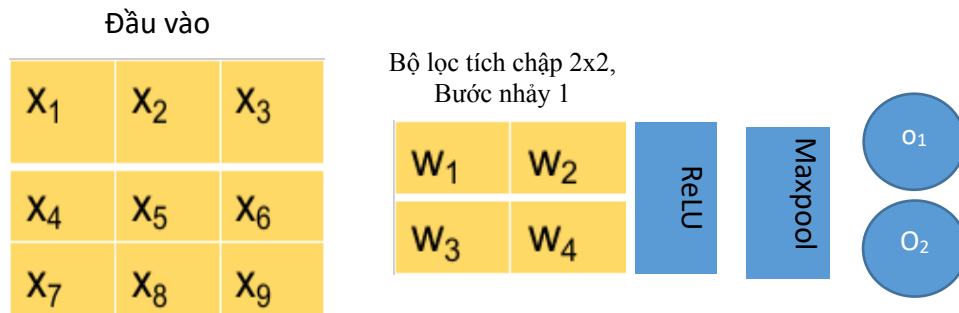
Cho mạng ANN phân loại chó và mèo như hình vẽ sau. Hàm kích hoạt tại các nút đều là hàm tanh. Hàm mất mát là hàm softmax, cross-entropy. Ban đầu các trọng số được khởi tạo là $w_1=1, w_2=2, w_3=3, w_4=4, w_5=5, w_6=6, w_7=3, w_8=4, w_9=5$ và $w_{10}=6$. Cho một mẫu mèo với $x_1=0,2, x_2=0,1$.



1. Tính toán giá trị đầu ra y và giá trị hàm mất mát với cơ chế lan truyền tiến? (1 điểm)
2. Với thuật toán lan truyền ngược, tính toán các giá trị gradient của hàm mất mát theo từng tham số $w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8, w_9$ và w_{10} ? (2 điểm)

Bài 3. (3 điểm) Mạng CNN

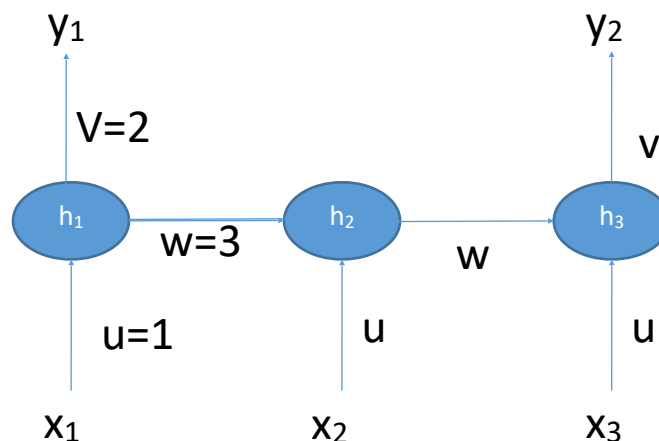
Cho mạng CNN dùng để hồi quy vị trí tâm quả bóng hình tròn có đường kính là 1 pixel trong một hình ảnh đen trắng 3x3. Chẳng hạn, một quả bóng ở vị trí $(o_1, o_2) = (1, 2)$ có hình ảnh $x_1=0, x_2=1, x_3=0, x_4=0, x_5=0, x_6=0, x_7=0, x_8=0, x_9=0$. Ở đây, ta có $o_1 = \text{max_pooling}(w_1, w_4), o_2 = \text{max_pooling}(w_2, w_3)$. Hàm mất mát được dùng là chuẩn 2 (L2). Các trọng số được khởi tạo là $w_1=1, w_2=2, w_3=3, w_4=4$, độ lệch cho tầng tích chập này là $b_1=0,5$.



1. Tính toán giá trị các đầu ra o_1, o_2 , giá trị hàm mất mát với cơ chế lan truyền tiến? (0,5 điểm)
2. Với thuật toán lan truyền ngược, hãy tính gradient của hàm mất mát so với từng trọng số w_1, w_2, w_3, w_4, b_1 ? (1,5 điểm)
3. Anh/chị hãy đánh giá hiệu quả của mô hình trên trong bài toán đặt ra. Hãy đề xuất cách thức cải tiến mô hình để đạt hiệu quả cao hơn? (1 điểm)

Bài 4. (3 điểm) Mạng RNN

Cho mạng nơ ron tích chập sau dùng để phân loại chuỗi đầu vào (x_1, x_2, x_3) là tăng dần hay giảm dần. Cho mẫu giảm dần $(x_1=3, x_2=2, x_3=1)$, hàm mất mát cross-entropy, hàm kích hoạt là hàm tanh.



1. Tính toán giá trị các đầu ra y_1 , y_2 , giá trị hàm mất mát với cơ chế lan truyền tiến? (0,5 điểm)

2. Với thuật toán lan truyền ngược, hãy tính gradient của hàm mất mát so với từng trọng số u, v, w ? (1,5 điểm)

3. Anh/chị hãy đánh giá hiệu quả của mô hình trên trong bài toán đặt ra. Hãy đề xuất cách thức cải tiến mô hình để đạt hiệu quả cao hơn? (1 điểm)