

Đề bài tập cuối kỳ học kỳ II (2020-2021)
Môn : Học Máy và Ứng dụng
Lớp Cao học
(Sinh viên được phép tham khảo slide môn học)

Đề gồm 3 trang

1. (0.5 điểm) Nêu điểm khác biệt giữa giải thuật *suy giảm độ dốc* (gradient descent) và giải thuật *suy giảm độ dốc tăng dần* (incremental gradient descent) khi huấn luyện một đơn vị perceptron tuyến tính.

2. (0.75 điểm) Khi xác định cấu hình của mạng nơ ron cho một bài toán phân lớp, căn cứ vào đâu để xác định số nút của tầng nhập và số nút của tầng xuất? Và bằng cách nào xác định số nút thích hợp cho tầng ẩn ?

3. (1.25 điểm)

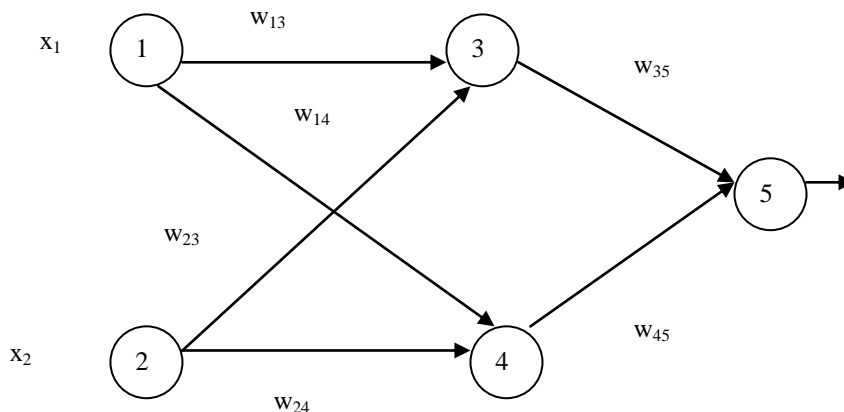
Điều chỉnh lại giải thuật *lan truyền ngược* để giải thuật này có thể làm việc được với các nút sử dụng hàm kích hoạt *tanh* thay vì hàm sigmoid. Hàm *tanh* được định nghĩa như sau:

$$\tanh(x) = (e^x - e^{-x}) / (e^x + e^{-x})$$

Tức là điều chỉnh lại qui tắc cập nhật trọng số tại tầng xuất và tầng ẩn. Lưu ý đạo hàm hàm *tanh* có tính chất: $\tanh'(x) = 1 - [\tanh(x)]^2$.

4. (1.75 điểm)

Cho mạng nơ ron có cấu hình như sau:



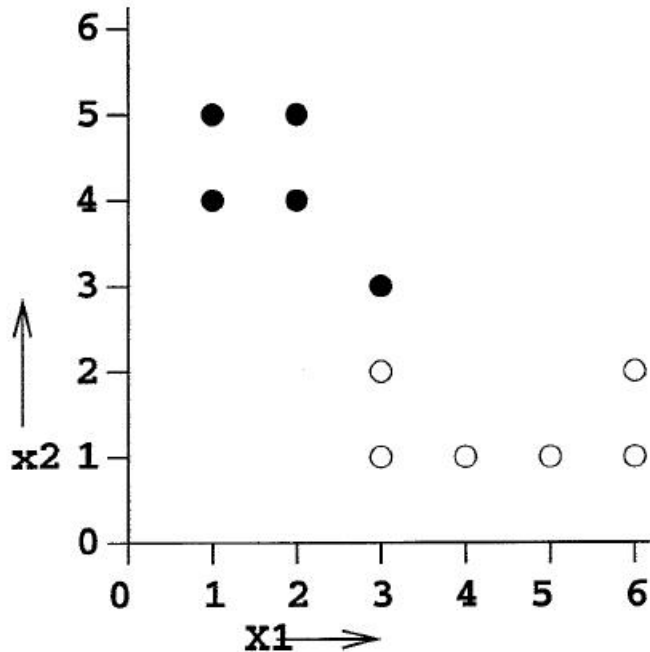
Cho các giá trị trọng số và các độ lệch (bias) khởi đầu như trong bảng sau:

w_{13}	w_{14}	w_{23}	w_{24}	w_{35}	w_{45}	w_{03}	w_{04}	w_{05}
0.2	-0.3	0.4	0.1	-0.3	-0.2	-0.4	0.2	0.1

Giả sử hệ số học $\eta = 0.8$, hàm sigmoid được dùng ở tầng ẩn và tầng xuất, và giải thuật lan truyền ngược được dùng để huấn luyện mạng nơ ron. Nếu mẫu huấn luyện $\langle x_1 = 0, x_2 = 1 \rangle$ và output $= 1$ được đưa vào giải thuật huấn luyện thì các trọng số và các độ lệch (w_{03}, w_{04}, w_{05}) sẽ được điều chỉnh lại như thế nào?

5. (1.75 điểm)

a) Cho một tập dữ liệu mà mỗi điểm dữ liệu có hai thuộc tính (x_1 and x_2) và một nhãn lớp như trong hình vẽ sau. Các điểm thuộc lớp dương ký hiệu bằng dấu chấm đậm và các điểm thuộc lớp âm ký hiệu bằng những hình tròn nhỏ. Giả sử chúng ta đang dùng SVM tuyến tính để phân lớp. Hãy vẽ 3 đường thẳng biểu diễn siêu phẳng tách có khoảng biên cực đại và hai siêu phẳng hỗ trợ. Tính khoảng biên cực đại. (1 điểm)



b. Hãy nêu công dụng của hàm kernel đối với máy véc tơ hỗ trợ trong trường hợp dữ liệu không khả tách một cách tuyến tính. (0.25 điểm)

c. Giải thích ý nghĩa của các biến bù ξ_i và thông số C trong máy véc tơ hỗ trợ với khoảng biên mềm (soft margin) (0.5 điểm)

6. (1.25 điểm) Cho tập mẫu gồm 14 mẫu hai chiều như sau:

$X_1 = (1, 1, 1)$, $X_2 = (1, 2, 1)$, $X_3 = (2, 1, 1)$, $X_4 = (2, 1.5, 1)$, $X_5 = (3, 2, 1)$,
 $X_6 = (4, 1.5, 2)$, $X_7 = (4, 2, 2)$, $X_8 = (5, 1.5, 2)$, $X_9 = (4.5, 2, 2)$, $X_{10} = (4, 4, 3)$,
 $X_{11} = (4.5, 4, 3)$, $X_{12} = (4.5, 5, 3)$, $X_{13} = (4, 5, 3)$, $X_{14} = (5, 5, 3)$,

Mỗi mẫu được biểu diễn bởi đặc trưng 1, đặc trưng 2 và nhãn lớp. Ta xây dựng một tổ hợp bộ phân lớp bằng kỹ thuật bagging như sau:

Khi tạo bộ phân lớp thứ nhất, bằng cách lấy mẫu, ta tạo ra tập huấn luyện cho bộ phân lớp này : $D1 = \{X_2, X_4, X_5, X_8, X_9, X_{12}, X_{13}, X_{14}\}$

Khi tạo bộ phân lớp thứ hai, bằng cách lấy mẫu, ta tạo ra tập huấn luyện cho bộ phân lớp này : $D2 = \{X_2, X_4, X_5, X_8, X_9, X_{10}, X_{11}\}$

Khi tạo bộ phân lớp thứ ba, bằng cách lấy mẫu, ta tạo ra tập huấn luyện cho bộ phân lớp này : $D3 = \{X_1, X_3, X_8, X_9, X_{12}, X_{13}, X_{14}\}$

Khi tạo bộ phân lớp thứ tư, bằng cách lấy mẫu, ta tạo ra tập huấn luyện cho bộ phân lớp này : $D4 = \{X_1, X_3, X_8, X_9, X_{10}, X_{11}\}$

Khi tạo bộ phân lớp thứ năm, bằng cách lấy mẫu, ta tạo ra tập huấn luyện cho bộ phân lớp này : $D5 = \{X_1, X_3, X_6, X_7, X_{12}, X_{13}, X_{14}\}$

Khi tạo bộ phân lớp thứ sáu, bằng cách lấy mẫu, ta tạo ra tập huấn luyện cho bộ phân lớp này : $D6 = \{X_1, X_3, X_6, X_7, X_{10}, X_{11}\}$

Các bộ phân lớp thành phần đều là bộ phân lớp 1-lần cận gần nhất.

Xét mẫu thử (3.5, 2.8). Bộ phân lớp tổ hợp sẽ quyết định mẫu thử thuộc lớp nào ?

7. (0.5 điểm) Nêu hai yếu tố ngẫu nhiên được đưa vào giải thuật tổ hợp phân lớp RandomForest.

8. (0.5 điểm) Nêu hai cách tiếp cận chính để giải quyết vấn đề mất cân bằng dữ liệu (imbalanced data) trong bài toán phân lớp.

9. (1.75 điểm)

a. Nêu công dụng chính của kiến trúc nhiều tầng trong mạng nơ ron học sâu.

(0.25 điểm)

b. Trong lãnh vực mạng nơ ron học sâu, *mô hình sinh* (generative model) và *mô hình phân biệt* (discriminative model) khác biệt nhau như thế nào?

(0.5 điểm)

c. Nêu tên hai giải thuật dùng trong hai giai đoạn của quá trình huấn luyện mạng Deep Belief Network; một giai đoạn là học không giám sát và một giai đoạn là học có giám sát.

(0.5 điểm)

d. Giải thích công dụng của tầng tích chập (convolution) và tầng gộp (pooling) trong mạng nơ ron tích chập.

(0.5 điểm)