

Họ và tên: Huỳnh Thiện Tùng

Mã số sinh viên: 19522492

Bài tập: BTLT1 - Softmax Regression mà có nhiều hơn 1 lớp thì có giải quyết được bài toán phi tuyến không?

Trả lời

Không thể. Để lý giải cho câu trả lời trên, ta nhắc lại một chút về kiến trúc mạng nơ ron (neural network). Trong mạng neural, từng lớp ẩn (hidden layer) sẽ có một hàm kích hoạt (activation function). Mục đích của hàm kích hoạt là đưa “tính năng” phi tuyến vào mạng. Tính phi tuyến tính là đầu ra (output) không thể tái tạo lại từ sự tổ hợp tuyến tính của đầu vào (input). Nếu không có hàm phi tuyến trong mạng neural thì dù cho có bao nhiêu lớp ẩn đi chăng nữa thì mạng neural sẽ hoạt động như một lớp perceptron mà chính nó mang một hàm tuyến tính (vì kết quả của layer tiếp theo được tính từ layer trước đó). Như vậy, dù cho deep neural network hay shallow neural network thì cũng chẳng khác gì nhau! Đối với câu hỏi này, thì không thể giải quyết được cho bài toán phi tuyến.

Một ví dụ điển hình trong NN là đầu vào là một phép toán tuyến tính ($\text{input} * \text{weight}$) nhưng trong thực tế các vấn đề mà chúng ta gặp phải đều không được tuyến tính như ta mong đợi. Input của mạng có thể bắt nguồn từ khoảng **-inf** đến **+inf**. Hàm kích hoạt phi tuyến giải quyết vấn đề này bằng cách scale miền giá trị về khoảng nhất định ví dụ như $\{0, 1\}$ hoặc $\{-1, 1\}$, ... Do đó vẫn cần một hàm phi tuyến dùng làm hàm kích hoạt cho mạng. Tóm lại là tính phi tuyến rất cần thiết trong mạng neural vì nó giải quyết được vấn đề tạo ra ranh giới (boundary decision line/plane) phi tuyến thông qua sự tổ hợp phi tuyến tính của $\text{input} * \text{weight}$