

### -Giải thích vì sao không dùng hàm MSE cho Logistics Regression

Ta có công thức MSE như sau:

$$\operatorname{argmin}(\theta) \frac{1}{2n} \sum_{i=1}^n (f(x_i) - y_i)^2$$

Khi sử dụng công thức này cho Logistics Regression, ta có vấn đề như sau

- $y_i$  là kết quả của bài toán classification nên giá trị trả về sẽ nằm trong khoảng  $[0;1]$
- Trong khi đó  $f(x_i) = \theta^T x$  có giá trị phụ thuộc vào  $x$  nên sẽ có khoảng giá trị  $[-\infty; +\infty]$ , khi đó giá trị của hàm Loss sẽ phụ thuộc nhiều vào  $f(x_i)$  (ví dụ  $f(x_i)$  có giá trị 100 còn  $y$  chỉ có giá trị 1) thì khi đó mô hình sẽ tập trung vào vấn đề tìm  $\theta$  để giảm giá trị  $f(x_i)$  xuống thay vì làm nhiệm vụ chính của nó là xấp xỉ 2 giá trị  $f(x_i)$  và  $y_i$
- Vậy nên cần một giải pháp ép được miền giá trị của  $f(x_i)$  về nhỏ hơn, lúc đó ta cần dùng hàm sigmoid với công thức  $\sigma(x) = \frac{1}{1+e^{-x}}$  với hình dạng đồ thị như bên dưới. Khi đó giá trị của  $f(x_i)$  sẽ được giới hạn về  $[0;1]$

