

Loss function for Logistic Regression

- Squared error (sai số bình phương) là hàm mất mát phù hợp cho **Linear Regression** vì nó tạo ra **hàm chi phí convex** (hình bát úp), giúp gradient descent hội tụ tốt. Tuy nhiên, khi áp dụng squared error cho **Logistic Regression**, hàm chi phí trở thành **không convex**, dẫn đến nhiều điểm cực tiểu cục bộ → gradient descent không còn hiệu quả.
- Hàm mất mát mới được định nghĩa dựa trên **log loss (logarithmic loss)**.
 - Nếu **y = 1**, loss = $\log(f(x))$.
 - Nếu **y = 0**, loss = $\log(1 - f(x))$.
- Hàm này có đặc điểm:
 - Khi dự đoán đúng ($f(x)$ gần y), loss nhỏ.
 - Khi dự đoán sai ($f(x)$ xa y), loss rất lớn → khuyến khích mô hình dự đoán chính xác.
- Ta có công thức tổng quát cho hàm loss:

$$Loss(f(x), y) = -[y \log(f(x)) + (1 - y) \log(1 - f(x))]$$

- Với m là số lượng mẫu huấn luyện, ta lấy **trung bình các mất mát**:

$$J(w, b) = -\frac{1}{m} \sum_{i=1}^m [y^{(i)} \log(f(x^{(i)})) + (1 - y^{(i)}) \log(1 - f(x^{(i)}))]$$