

# Logistic Regression from the Grounth Up

Nhóm MLOps

Ngày 18 tháng 10 năm 2025

Nội dung về Logistic Regression sẽ chia thành 2 phần chính, phần 1 giới thiệu về cả Linear Regression và tại sao lại cần Logistic Regression cũng như cách xây dựng nó. Phần 2 tập trung giới thiệu sâu hơn về việc tại sao lại sử dụng những công thức này và ý nghĩa của chúng.

- **Phần 1: Tổng quan về Logistic Regression và Linear Regression và Logistic Regression**
- **Phần 2: Ma trận: Nhấn mạnh vào Logistic Regression**
- **Phần 3: Tài liệu**

## Phần 1: Tổng quan về Linear Regression và Logistic Regression

### 1.1 Phần 1.1: Về Linear Regression

Mục đích của Linear Regression là tìm ra mô hình tốt nhất để mô tả mối quan hệ giữa các biến đầu vào  $x_i$  và biến đầu ra  $y_i$  (liên tục).

- Linear (tuyến tính) nghĩa là mối quan hệ giữa các biến đầu vào và biến đầu ra là một đường thẳng.
- Regression (hồi quy) là một phương pháp thống kê để mô tả mối quan hệ giữa các biến đầu vào và biến đầu ra.

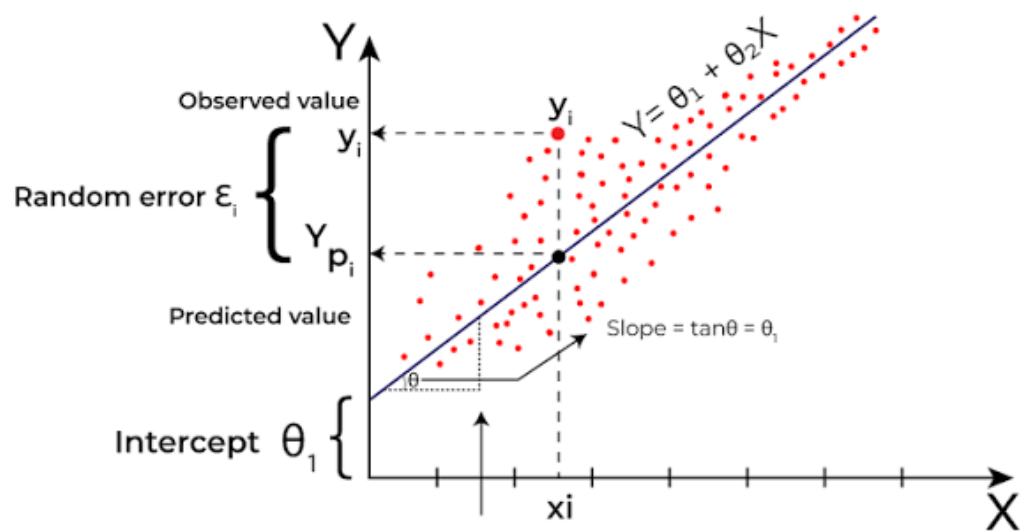


Figure 1: Linear Regression

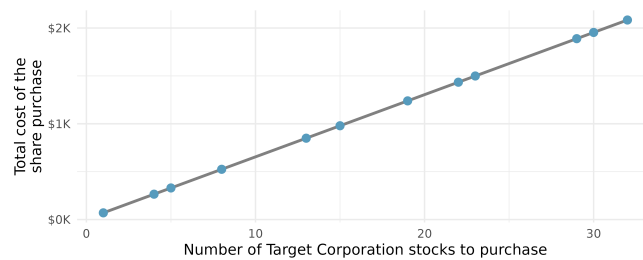


Figure 2: Trong iu kin hoìn ho, Linear Regression s khp vi d liu ãng 1 ng thng

Tuy nhiên vi d liu khũng tuyền tòngh vi liỏn tc, thờ Linear Regression khũng th khp c. Vy lỳm th nộ chn c 1 hỳm Loss giúp phỏn loi nhiu lp ? u tiỏn mỏnh cn hũu iu kin ca 1 hỳm Loss lỳ gờ trc.

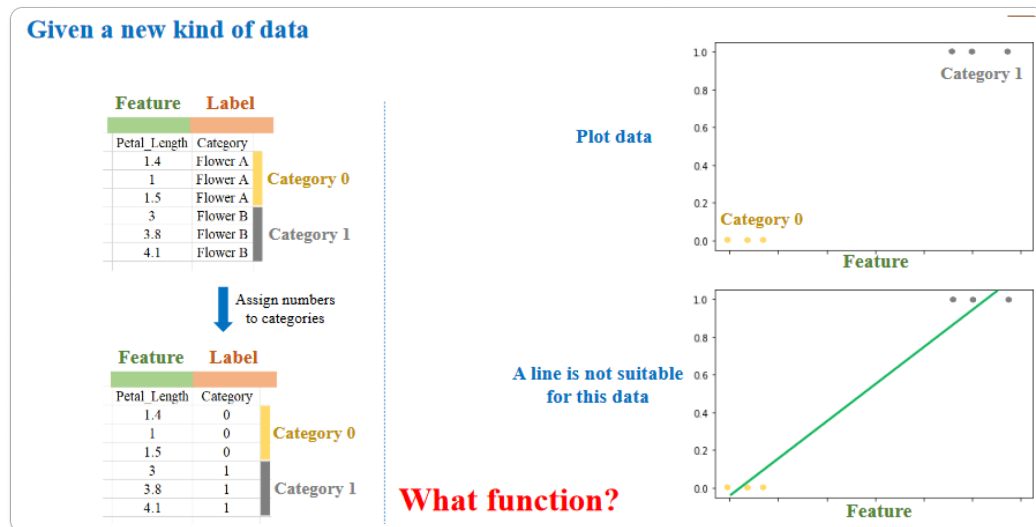


Figure 3: Non-Linear Data

- gii thorch cng thc qua 1 v d t u n cui. v d + nh

## 1.2 Phn 1.2: Logistic Regression

- Li n kt vi Linear Regression vi vn Logistic Regression gii quy t - gii thorch thut ng - gii thorch cng thc qua 1 v d t u n cui 1. Initialize Parameters 2. Calculating  $y \rightarrow$  Calculate Sigmoid  $z$  3. Calculate Loss base on  $\hat{y}$  and  $y$  4. Calculate Parameters Gradient 5. Update Parameters Gradient ( $w$  and  $b$ )

- gii thorch y ngha cng thc - code li s dng v d.

## 1.3 Phn 1.3: X y dng h m Loss t Suy Lun

## 1.4 Nhc im ca MSE

+ T orch tin n 0. + Non-Convex  $\rightarrow$  kh u hi t. (Gii thorch s u hi sau khi x y dng xong BCE)

## Logistic Regression-MSE

❖ Result

**Model and Loss**

$$z = \theta^T x = x^T \theta$$

$$\hat{y} = \sigma(z) = \frac{1}{1 + e^{-z}}$$

$$L = (\hat{y} - y)^2$$

$$\frac{\partial L}{\partial \theta_i} = 2x_i(\hat{y} - y)\hat{y}(1 - \hat{y})$$

**Done?**

Feature	Label
Petal.Length	Category
1.4	0
1	0
1.5	0
3	1
3.8	1
4.1	1

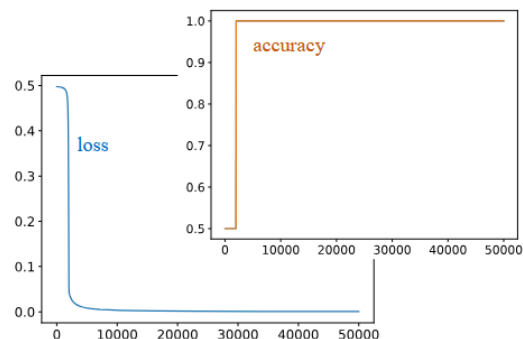


Figure 4: Problem with Non-Convex Loss Function

## 1.5 Lựa chọn vị Xảy dạng hàm Loss cho bài toán phân loại

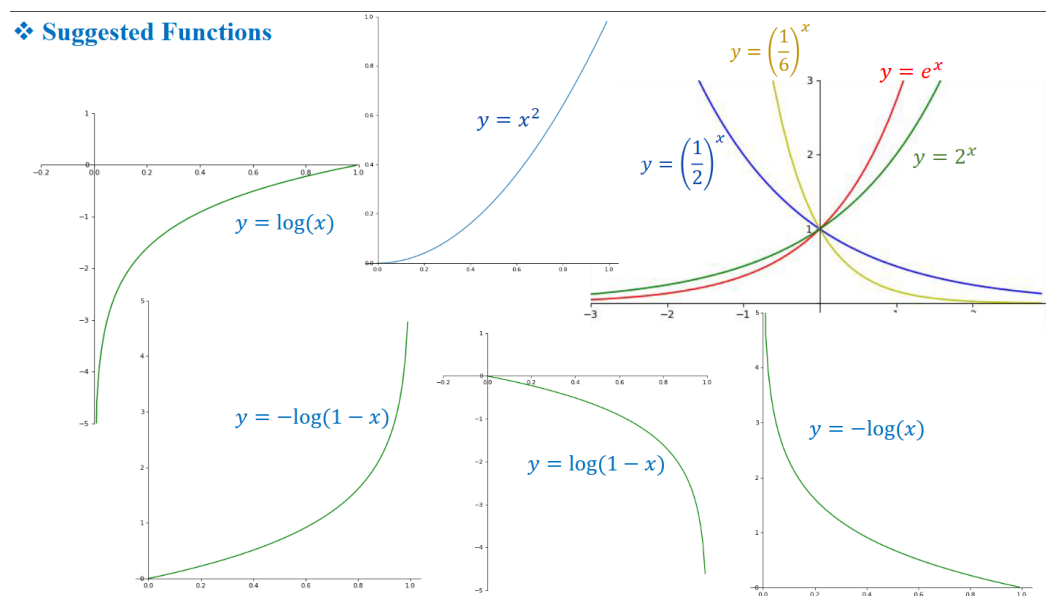


Figure 5: Suggested Loss Function

### ❖ Loss function

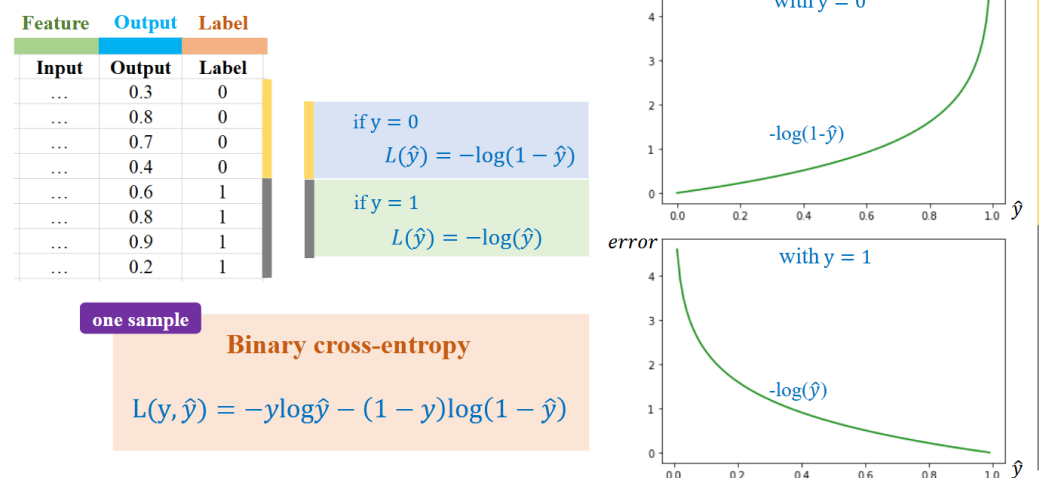


Figure 6: Binary Cross Entropy Loss

## Construct loss

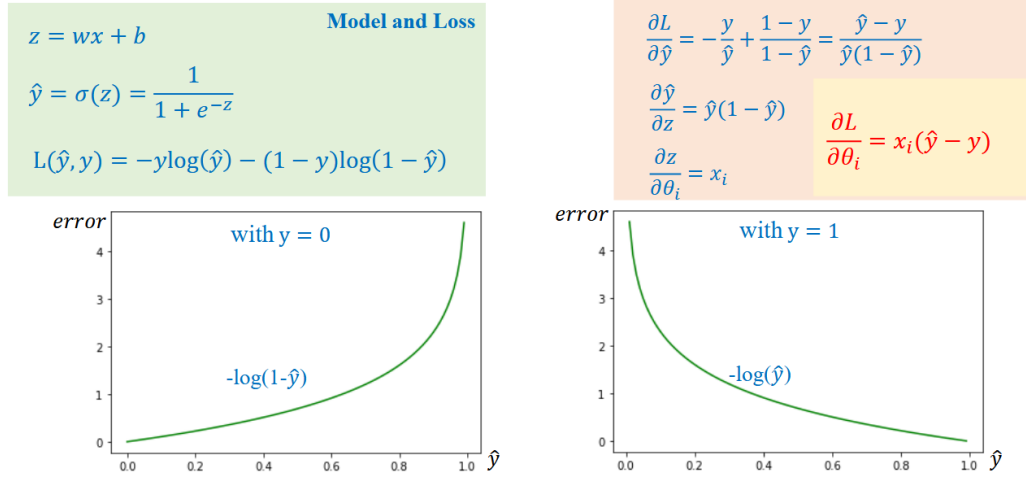


Figure 7: BCE Summarize

Explain Convexity using that Coursera Linear Algebra Course

## 1.6 Chng Minh Convex ( li) ca MSE vj BCE

Chng Minh Convex ( li) ca MSE

## Mean Squared Error

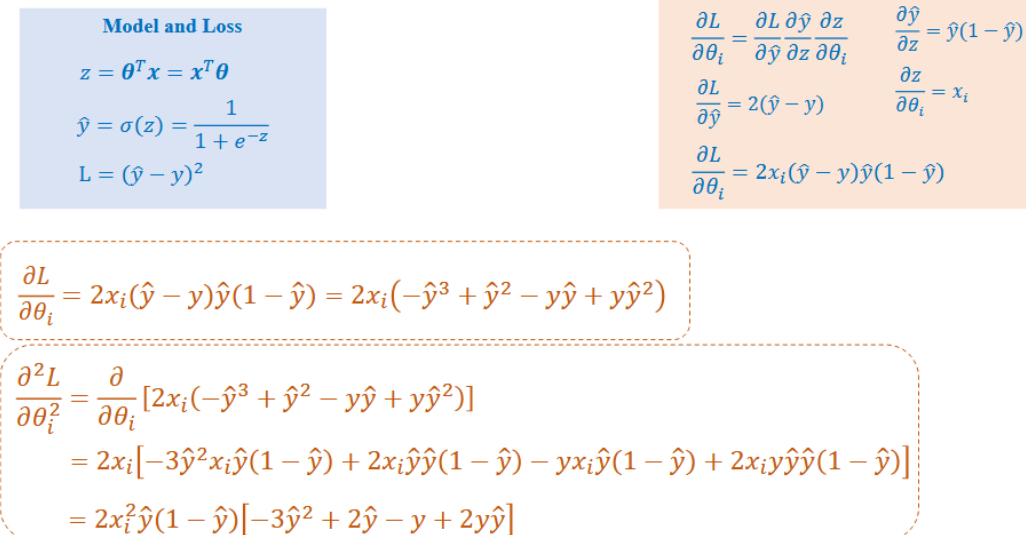


Figure 8: MSE vs BCE

## Mean Squared Error

$$\frac{\partial^2 L}{\partial \theta_i^2} = 2x_i^2 \hat{y}(1 - \hat{y})[-3\hat{y}^2 + 2\hat{y} - y + 2y\hat{y}]$$

$$x_i^2 \geq 0$$

$$\hat{y}(1 - \hat{y}) \in \left[0, \frac{1}{4}\right]$$

$$y = 0$$

$$f(\hat{y}) = -3\hat{y}^2 + 2\hat{y}$$

$$y = 1$$

$$f(\hat{y}) = -3\hat{y}^2 + 4\hat{y} - 1$$

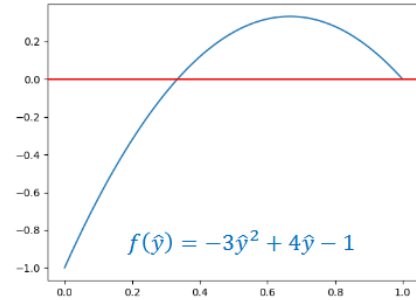
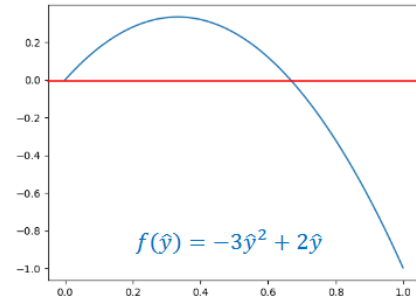
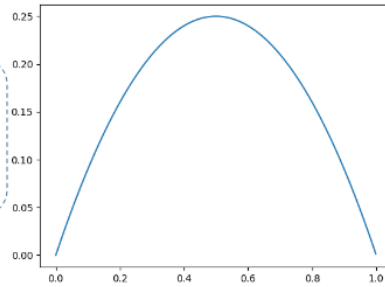
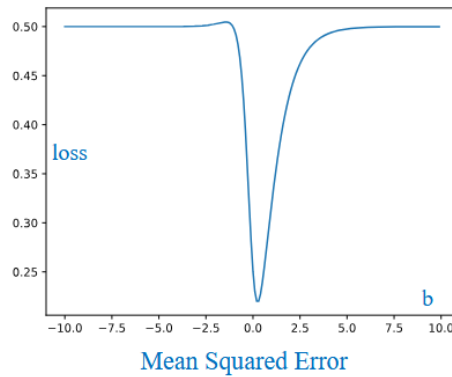


Figure 9: MSE vs BCE



### Model and Loss

$$z = \theta^T x = x^T \theta$$

$$\hat{y} = \sigma(z) = \frac{1}{1 + e^{-z}}$$

$$L = (\hat{y} - y)^2$$

$$\frac{\partial L}{\partial \theta_i} = 2x_i(\hat{y} - y)\hat{y}(1 - \hat{y})$$

Figure 10: MSE vs BCE

Chng Minh Convex ( li) ca BCE

# Binary Cross-entropy

## Convex function

$z = \theta^T x$  **Model and Loss**

$$\hat{y} = \sigma(z) = \frac{1}{1 + e^{-z}}$$

$$L = -y \log(\hat{y}) - (1 - y) \log(1 - \hat{y})$$

simplified version

$$\frac{\partial L}{\partial \theta_i} = x_i(\hat{y} - y)$$

$$\frac{\partial^2 L}{\partial \theta_i^2} = \frac{\partial}{\partial \theta_i} [x_i(\hat{y} - y)] = x_i^2(\hat{y} - \hat{y}^2) \geq 0$$

$$x_i^2 \geq 0 \quad \hat{y} - \hat{y}^2 \in \left[0, \frac{1}{4}\right]$$

$$\frac{\partial L}{\partial \theta_i} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial z} \frac{\partial z}{\partial \theta_i} \quad \text{Derivative}$$

$$\frac{\partial L}{\partial \hat{y}} = -\frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}} = \frac{\hat{y} - y}{\hat{y}(1-\hat{y})}$$

$$\frac{\partial \hat{y}}{\partial z} = \hat{y}(1-\hat{y})$$

$$\frac{\partial z}{\partial \theta_i} = x_i \quad \frac{\partial L}{\partial \theta_i} = x_i(\hat{y} - y)$$

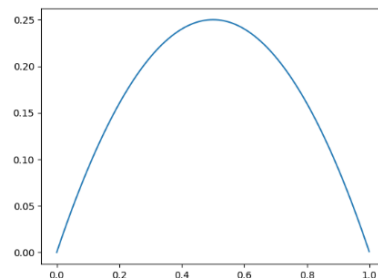


Figure 11: BCE Convex

## 1.7 So sánh MSE và BCE

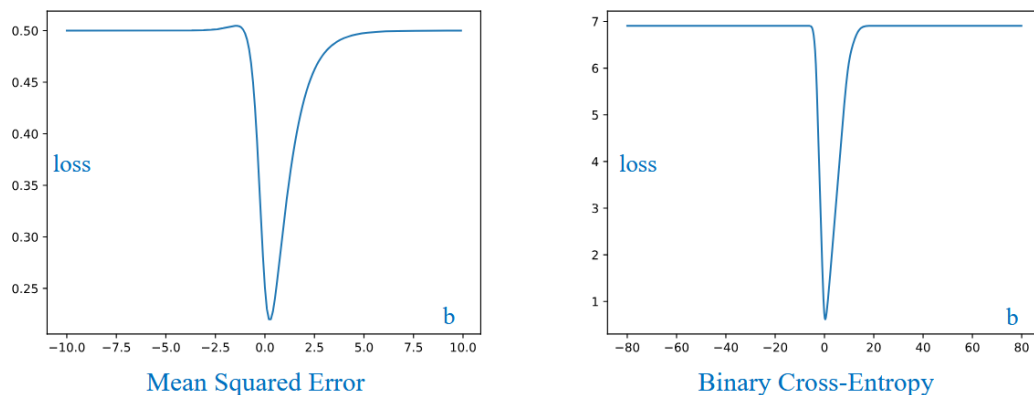


Figure 12: MSE vs BCE

**MSE:** gentle curve, Phù hợp cho các dữ liệu sai. **BCE:** shallow curve, Phù hợp cho các dữ liệu sai. Vì BCE là lựa chọn tốt nhất cho Loss cho Logistic Regression.

## 1.8 Phần 1.4: Xây dựng hàm Loss và Xác suất Thống Kê

## 1.9 Giới thiệu Binary-Cross Entropy và góc nhìn của Xác suất Thống Kê

Entropy là gì? Entropy measure information, blah blah

Entropy thì hình gì trong Logistic Regression ?

## Phn 2: M Rng: Nhón ma trn vi Logistic Regression

- Gii thòch 1 s lý thuyt i s tuyền tòngh trong nhón ma trn. + Nhón nghch o - Vò d vi ma trn X vi w.  
Nhón ma trn. -> tng tòngh tng tòc. Gi ý thay bin ng c hc qua tng tòc.

### 2.1 Convexity of BCE for Multiple-Variable using Hessian Matrix

## Phn 3: Ti sao

3.1 Ti sao li nhón nghch o ?

3.2 Ti sao li s dng hịm tng trng Sigmoid ?

3.3 Ti sao li s dng  $\ln()$  thay vớ  $\log()$

3.4 Mi liỏn h gia  $\ln()$  vi Sigmoid ?

3.5 Ti sao li s dng loglikelihood thay vớ probability

3.6 Convex lị gớ ?

3.7 Convex trong Logistic Regression cho tòngh toỏn ma trn