

Logistic Regression from the Ground Up

Nhóm MLOps

Ngày 18 tháng 10 năm 2025

Nội dung về Logistic Regression có chia thành 2 phần chính, phần 1 giới thiệu về Linear Regression và sau đó là Logistic Regression cùng nhau xóa dung nhan. Phần 2 tập trung giới thiệu về các vấn đề liên quan đến việc áp dụng Logistic Regression.

- Phân 1: T Linear Regression và Logistic Regression
- Phân 2: M Rng: Nhóm ma tròn và Logistic Regression
- Phân 3: Ti Sao

Phân 1: T Linear Regression và Logistic Regression

1.1 Phân 1.1: Vai trò của Linear Regression

Ý tưởng của Linear Regression là để tìm một mô hình tuyến tính để mô tả mối quan hệ giữa biến độc lập x_i và biến phụ thuộc y_i (i là số thứ tự).

- Linear (tuy nhiên) nghĩa là mô hình quan hệ là một đường thẳng.
- Regression (hiệu ứng) là một phương pháp dùng để tìm mô hình quan hệ là một đường thẳng.

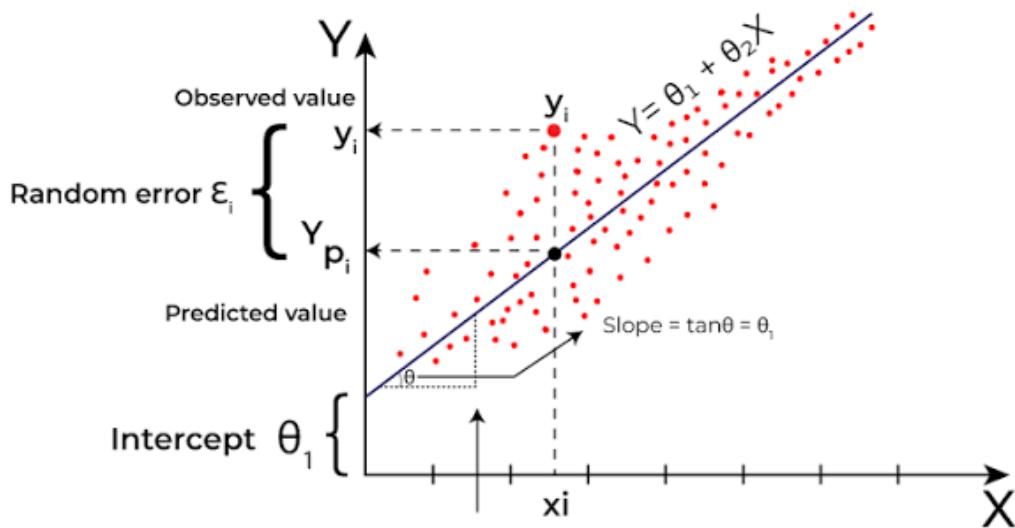


Figure 1: Linear Regression

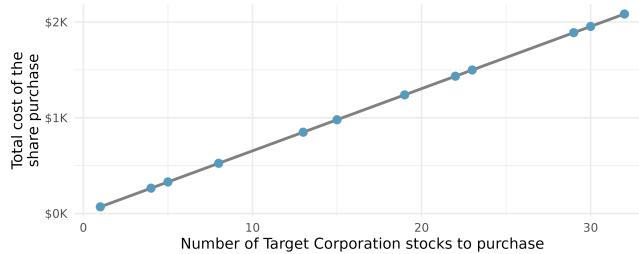


Figure 2: Trong iu kin hojn ho, Linear Regression s khp vi d liu ũng 1 ng thng

Tuy nhn vi d liu khng tuyn tnhn v lnn tc, tho Linear Regression khng th khp c. Vy ljm th nio chn c 1 hjm Loss giüp phón loi nhu lp ? u tihn mnh cn hiu iu kin ca 1 hjm Loss l gø trc.

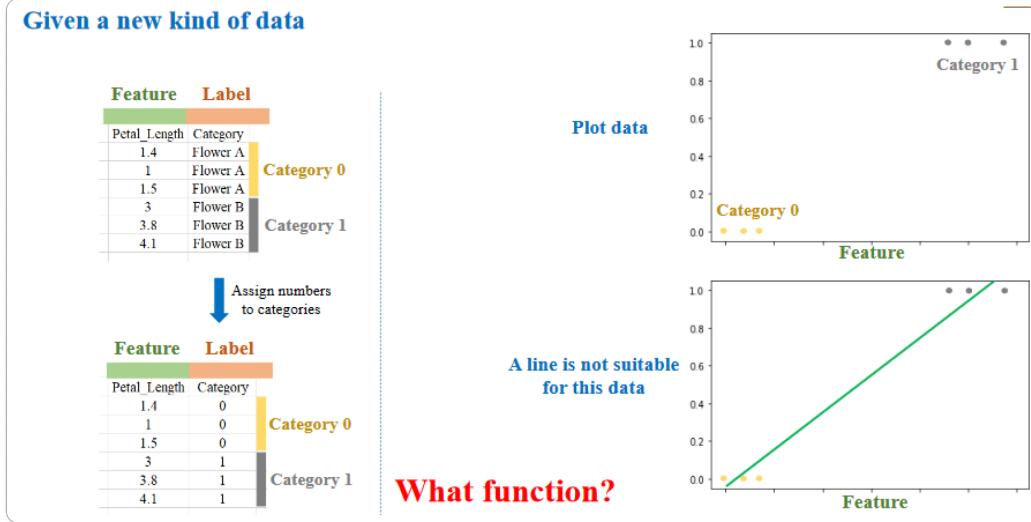


Figure 3: Non-Linear Data

- gii thòch cùng thc qua 1 vò d t u n cui. vò d + nh

1.2 Phn 1.2: Logistic Regression

- Liển kt vi Linear Regression vị vn Logistic Regression gii quyết - gii thòch thut ng - gii thòch cùng thc qua 1 vò d t u n cui 1. Initialize Parameters 2. Calculating y -> Calculate Sigmoid z 3. Calculate Loss base on \hat{y} and y 4. Calculate Parameters Gradient 5. Update Parameters Gradient (w and b)
- gii thòch y ngha cùng thc - code li s dng vò d.

1.3 Phn 1.3: Xóy dng hım Loss t Suy Lun

1.4 Nhć im ca MSE

+ Tòch tin n 0. + Non-Convex -> khú hi t. (Gii thòch sóu hi sau khi xóy dng xong BCE)

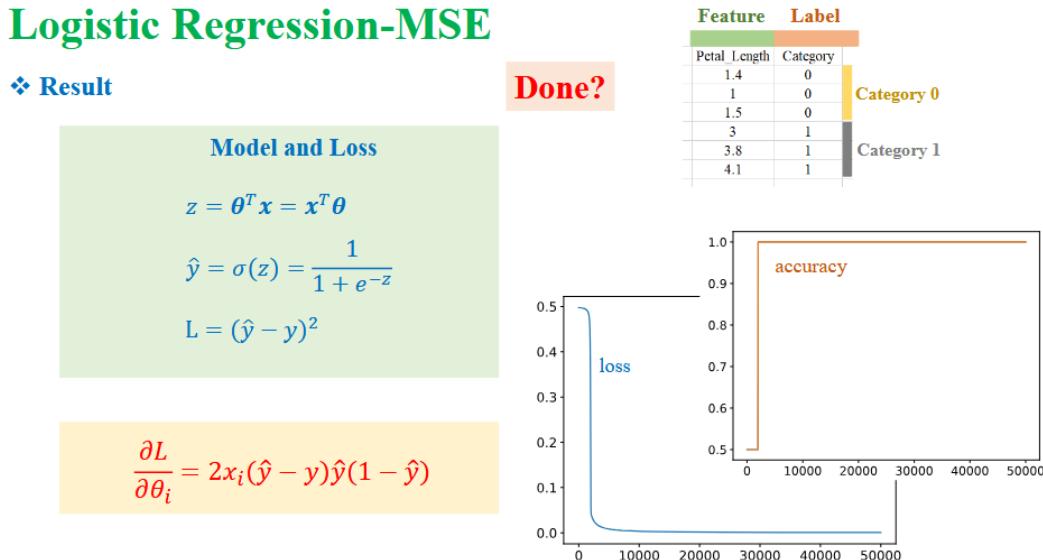


Figure 4: Problem with Non-Convex Loss Function

1.5 La chn vị Xóy dng hjm Loss cho bii toòn phón loi

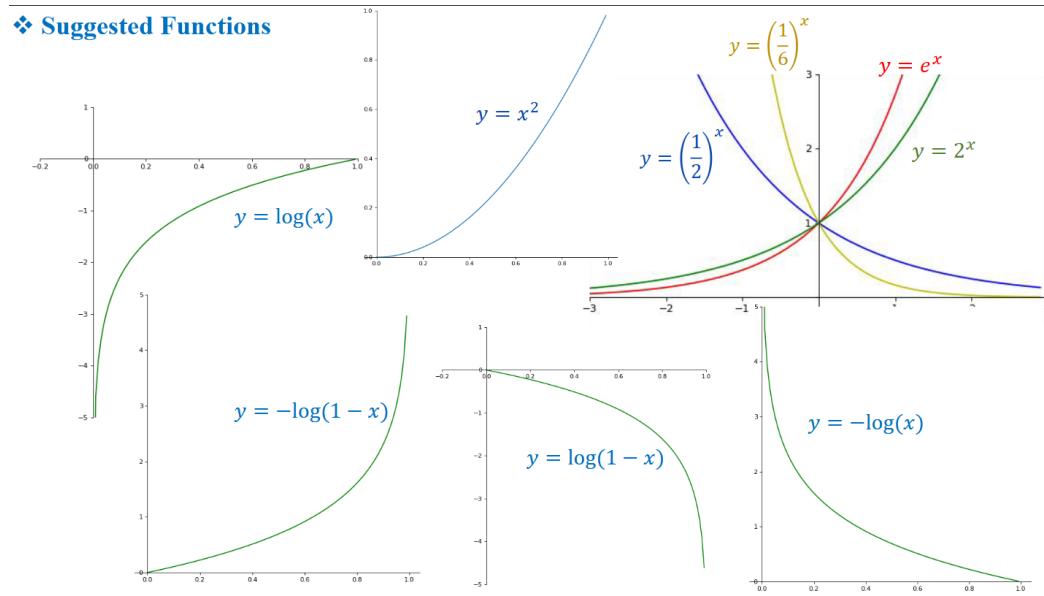


Figure 5: Suggesst Loss Function

❖ Loss function

Feature	Output	Label
Input	Output	Label
...	0.3	0
...	0.8	0
...	0.7	0
...	0.4	0
...	0.6	1
...	0.8	1
...	0.9	1
...	0.2	1

one sample

Binary cross-entropy

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

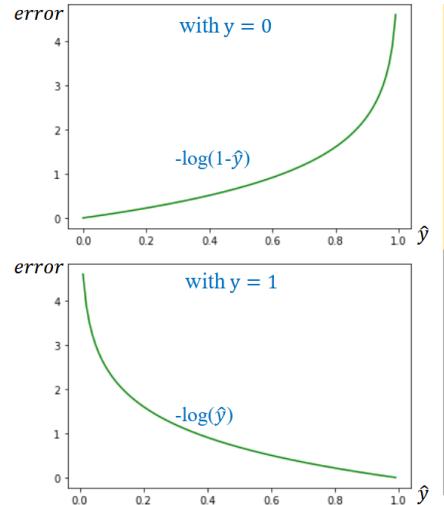


Figure 6: Binary Cross Entropy Loss

Construct loss

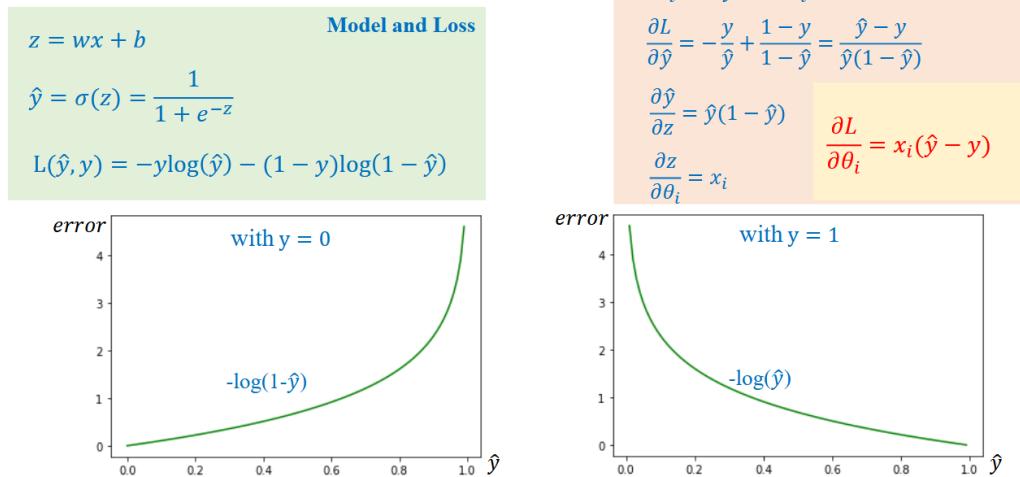


Figure 7: BCE Summarize

Explain Convexity using that Coursera Linear Algebra Course

1.6 Chng Minh Convex (li) ca MSE vi BCE

Chng Minh Convex (li) ca MSE

Mean Squared Error



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

$$\begin{aligned} \frac{\partial^2 L}{\partial \theta_i^2} &= \frac{\partial}{\partial \theta_i} [2x_i(-\hat{y}^3 + \hat{y}^2 - y\hat{y} + y\hat{y}^2)] \\ &= 2x_i[-3\hat{y}^2x_i\hat{y}(1 - \hat{y}) + 2x_i\hat{y}\hat{y}(1 - \hat{y}) - yx_i\hat{y}(1 - \hat{y}) + 2x_iy\hat{y}\hat{y}(1 - \hat{y})] \\ &= 2x_i^2\hat{y}(1 - \hat{y})[-3\hat{y}^2 + 2\hat{y} - y + 2y\hat{y}] \end{aligned}$$

Figure 8: MSE vs BCE

Mean Squared Error

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

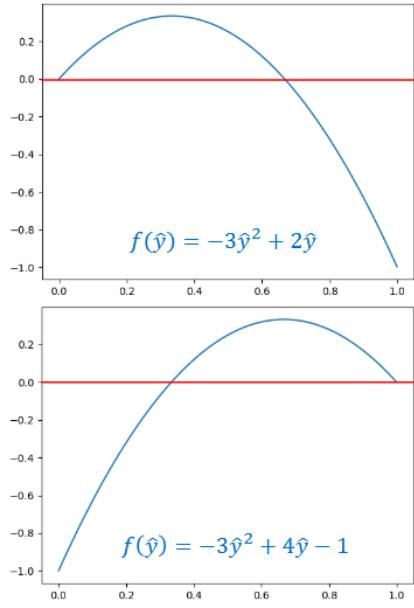
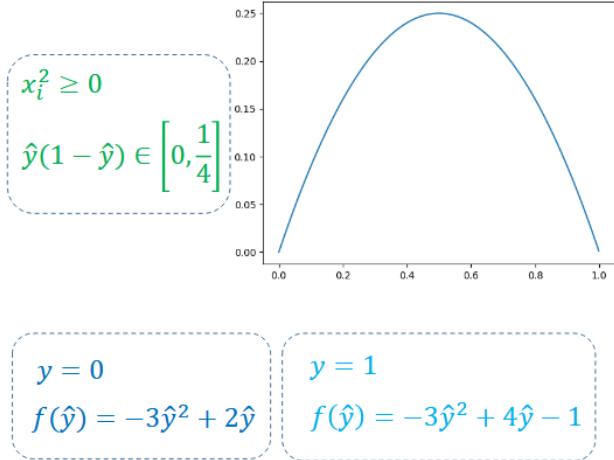
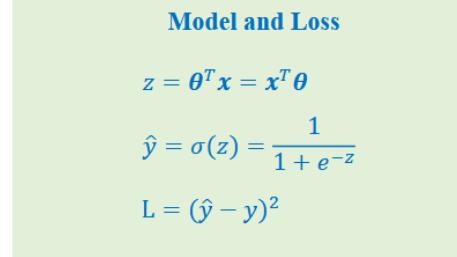
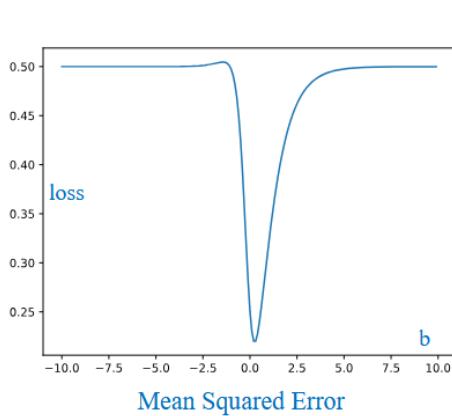


Figure 9: MSE vs BCE



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

Figure 10: MSE vs BCE

Chng Minh Convex (li) ca BCE

Binary Cross-entropy

Convex function

$$z = \theta^T x$$

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

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

Model and Loss

simplified version

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

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

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

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

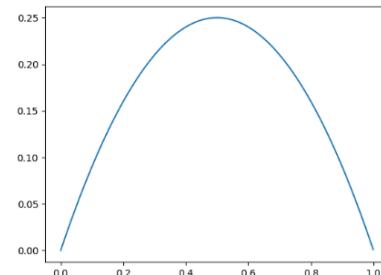


Figure 11: BCE Convex

1.7 So sánh MSE và BCE

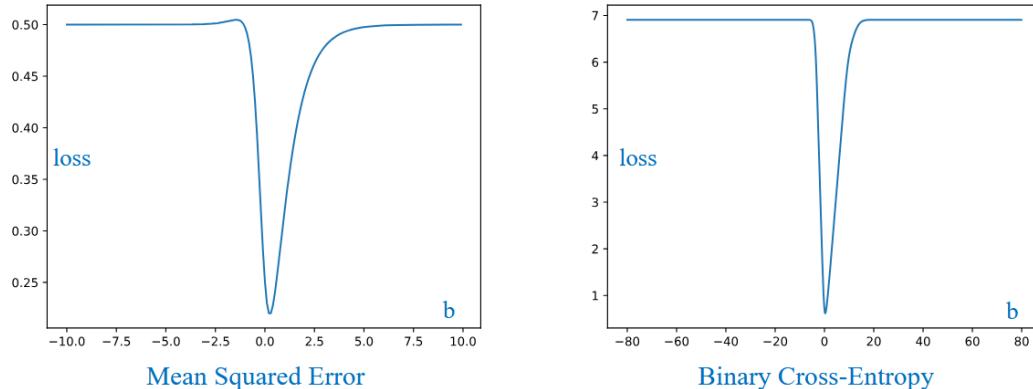


Figure 12: MSE vs BCE

MSE: gentle curve, Phù hợp với các lỗi nhỏ. **BCE:** shallow curve, Phù hợp với các lỗi lớn. Về BCE là cách đơn giản hóa Loss cho Logistic Regression.

1.8 Phân 1.4: Xây dựng hàm Loss t Xòc Sut Thng Kô

1.9 Giải thích Binary-Cross Entropy t giá trị nhọn ca Xòc Sut Thng Kô

Entropy là gì? Entropy measure information, blah blah

Entropy tham chiếu trong Logistic Regression ?

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

- Gii thòch 1 s lỳ thuỷt i s tuyn tònh trong nhón ma trn. + Nhón nghịch o - Vờ d vi ma trn X vị w. Nhón ma trn. -> tng tònh tng tóc. Gi y thay bin ng c hc qua tng tóc.

2.1 Convexity of BCE for Multiple-Variable using Hessian Matrix

Phn 3: Ti sao

- 3.1 Ti sao li nhón nghịch o ?
- 3.2 Ti sao li s dng hjm tng trng Sigmoid ?
- 3.3 Ti sao li s dng ln() thay vø log()
- 3.4 Mi liён h gia ln() vị Sigmoid ?
- 3.5 Ti sao li s dng loglikelihood thay vø probability
- 3.6 Convex lị gợ ?
- 3.7 Convex trong Logistic Regression cho tònh toòn ma trn