C in logistic Regression

11 August 2025 21:24

C = inverse of $\lambda = \dfrac{1}{\lambda}$ → regularization strength

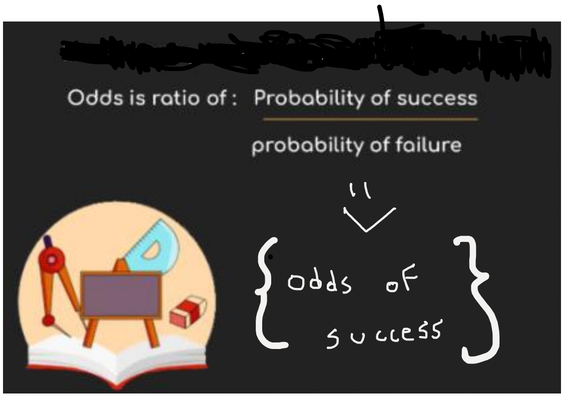
High C ⇒ less regularization ⇒ More overfitting

low C ⇒ More reg ⇒ More Underfitting

Balancing C ⇒ perfect fit

Odds Interpretation

10 August 2025 20:53



$$\text{odds of Failure} = \frac{\text{probability of Failure}}{\text{probability of success}}$$

{ If probability of Passing an exam = 0.8 }

$P(\text{Failure}) = 1 - 0.8 = 0.2$

$$\text{Odds of passing} = \frac{0.8}{0.2} = 4$$

$\text{odds of Failing} = 0.2 / 0.8 = 0.25$

$$\left\{ \text{Odds} = \frac{P}{1-P} \right\}$$

Log Odds

10 August 2025 21:04

$$\text{Odds} = \frac{P}{1-P}$$

Sigmoid ⇒ probabilities. Converts z into probability

$$\left\{ P = \frac{1}{1+e^{-z}} \right\} \Rightarrow 1+e^{-z} = \frac{1}{P}$$

$$e^{-z} = \frac{1}{P} - \frac{1}{1} \Rightarrow e^{-z} = \frac{1-P}{P}$$

$$\left. e^{z} = \frac{P}{1-P} \right\}$$

Negative to positive power converted by taking reciprocal

Take $\log_e$ on both sides

$$\log_e e^{z} = \log_e (P/1-P)$$

$$z \log_e e = \log_e (P/1-P)$$

$$z = \log_e (P/1-P) \Rightarrow P/1-P = \text{odds}!!$$

$$z = \text{Log odds}!!$$

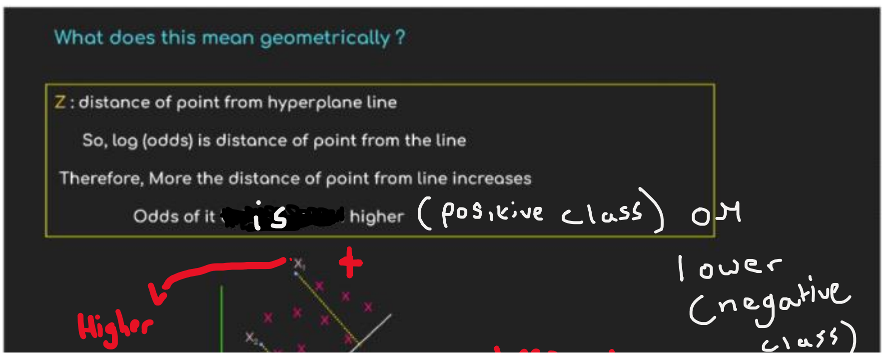
} Higher z = more odds = Higher P of positive class

$$\log_e e^{z}$$

$$= z \times \cancel{\log_e e}$$

$$= z \times 1$$

$$= z$$



{ Logistic Regression is all Modelling odds !! }

Continuous value

interview !!

odds

0.0001 or 9999

$P(s) = 0.99999$

$$\text{odds} = \frac{0.999}{0.001} = \left\{ \begin{array}{c} 99999 \\ \\ 0.00001 \end{array} \right\}$$

Difficult to
Model

$\log 99999 \} 7$

$\log 0.00001 \} -7$

$\} \Rightarrow$ Easier to Model

quiz

11 August 2025 21:52

$\rightarrow z$

How are log odds transformed into probabilities in logistic regression?

27 users have participated

A	By applying the sigmoid function	70%
B	By taking the exponential function	7%
C	By dividing by the odds ratio	22%
D	By subtracting the intercept term	0%

z is transformed into prob by passing through the sigmoid

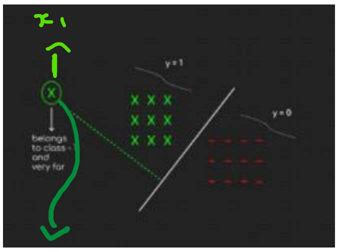
log odds = Z

{ Impact of outliers }

10 August 2025 21:18

In ML classification,

Outliers are not an issue when they are on the correct side.

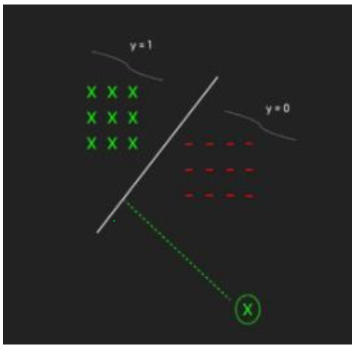


$y_{actual} = 1$

$y_{predicted} = 0.99$ ML Model predicted correctly!

⇐ ⇒ x_1 is far away from the boundary when compared to other points. However, it { still belongs to the correct class } ⇒ Means contributes positively to likelihood and hence, value of { loss function reduces } ↳ Good for us!!

in logistic { loss = negative log likelihood }



Outliers are a big problem when points lie far away from the boundary but on the wrong side.

$y_{actual} = 1$

$y_{predicted} = 0.01$

⇒ In this case, y_i (actual) = 1

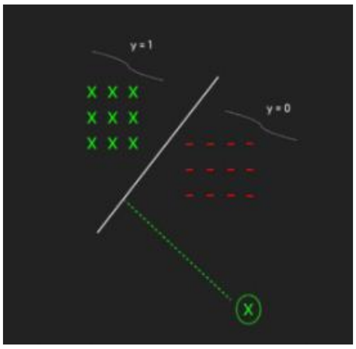
$\hat{y}_i$ (predicted probability) is very low (since it is far away from boundary on opposite side).

This means likelihood for that point will be very low ⇒ loss function increases!

Why is it such a big problem?

Because in the next iteration, gradient descent will try to fix the high value of loss.

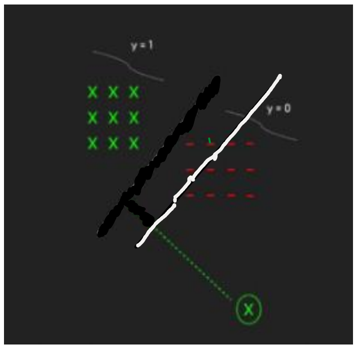
Result of GD|MLE at iteration t



Accuracy = $\frac{25}{26}$ = 96%

⇒ Boundary shifts because GD wants to reduce overall loss by reducing loss of that one point !

Result of GD|MLE at iteration t+1



⇒ G.D tried to over correct

Observe how the shift has caused more problems than before!

Accuracy = 22/26 = 88%

Solution → Remove outliers from data ! } ⇒ for logistic regression

Identify outliers using boxplots or remove based on percentiles

If you don't want to remove

↳ Treat them ⇒ Cap them

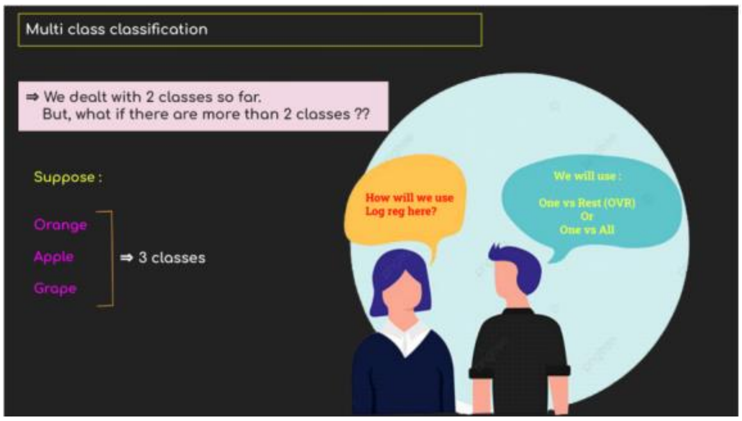
How do outliers affect the classification boundaries in logistic regression?

30 users have participated

A	Outliers shift the classification boundaries closer to the outlier values	83%
B	Outliers have no effect on the classification boundaries	3%
C	Outliers widen the gap between the classification boundaries	7%
D	Outliers make the classification boundaries more sensitive to minor changes	7%

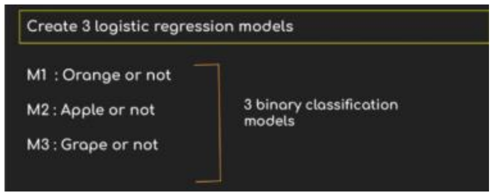
Multi Class Classification using Logistic Regression

10 August 2025 21:47



Originally logistic regression is for binary classification

0 / 1



Can I extend the concept of Logistic Regression for Multi-class

A v Non A
(1) (0)

id	x1	x2	y_A_vs_rest
1	2.1	0.5	1
2	1.3	1.1	0
3	0.4	2.3	0
4	2.2	0.3	1
5	1.4	0.9	0
6	0.1	1.9	0
7	2.0	0.6	1
8	1.0	1.5	0
9	0.3	2.0	0

→ refers to A

} → refers to some other class

Original dataset (example)

id	x1	x2	class
1	2.1	0.5	A
2	1.3	1.1	B
3	0.4	2.3	C
4	2.2	0.3	A
5	1.4	0.9	B
6	0.1	1.9	C
7	2.0	0.6	A
8	1.0	1.5	B
9	0.3	2.0	C

B(1) v NonB (0)

id	x1	x2	y_B_vs_rest
1	2.1	0.5	0
2	1.3	1.1	1
3	0.4	2.3	0
4	2.2	0.3	0
5	1.4	0.9	1
6	0.1	1.9	0
7	2.0	0.6	0
8	1.0	1.5	1
9	0.3	2.0	0

C(1) vs Non C (0)

id	x1	x2	y_C_vs_rest
1	2.1	0.5	0
2	1.3	1.1	0
3	0.4	2.3	1
4	2.2	0.3	0
5	1.4	0.9	0
6	0.1	1.9	1
7	2.0	0.6	0
8	1.0	1.5	0
9	0.3	2.0	1

Split original dataset into 3 datasets

Model Each dataset

Model 1 gives probability of A

Model 2 gives probability of B

Model 3 gives probability of C

How to classify?

For each point, there will be 3 predictions from 3 models. For example, take a point x_1:

$$\left.\begin{aligned} P(A) &= 0.9 \\ P(B) &= 0.2 \\ P(C) &= 0.3 \end{aligned}\right\} \Rightarrow$$ consider highest probability as predicted class

x_1 belongs to class A!

Original dataset (example)

id	x1	x2
1	2.1	0.5
2	1.3	1.1
3	0.4	2.3
4	2.2	0.3
5	1.4	0.9
6	0.1	1.9
7	2.0	0.6
8	1.0	1.5
9	0.3	2.0

A	B	C
Model 1	Model 2	Model 3
0.8	0.4	0.2

Belongs to A

$$\left\{\begin{aligned} P(A) &= 0.2 \\ P(B) &= 0.4 \\ P(C)\ &0.3 \end{aligned}\right\} \Rightarrow B$$

{One Vs Rest}

11 August 2025 22:52

How is the loss function typically defined in multi-class logistic regression?

22 users have participated

A	Cross-entropy loss	50%
B	Mean squared error (MSE)	41%
C	Mean absolute error (MAE)	0%
D	Hinge loss	9%

log loss = negative log likelihood