Logistic Regression Quiz

Total points 6/20

Email address *

skkumarlokam@gmail.com

Name *

Siva Lokam

✓ 1. A scenario where you have been given a fair coin and you want to find out the odds of getting heads. Which of the following option is true for such a case? 1/1

- ○ odds will be 0
- ○ odds will be 0.5
- ◉ odds will be 1 ✓
- ○ None of these

✓ 2. One of the good way to analyse performance of Logistic Regression is AIC, which is similar to R-Squared in Linear Regression. Which of the following is true for a good model ? 1/1

- ◉ A model with minimum AIC value ✓
- ○ A model with maximum AIC value
- ○ Both but depends on the situation
- ○ None of these

✓ 3. In Linear Regression, we train the model to get the optimium coefficient . In the similar manner, while training logistic regression what are we optimising ? 1/1

- ○ RMSE
- ◉ Log Loss ✓
- ○ Deviance
- ○ None of the above

✗ 4. Parameters in logistic regression can be regularized using?? 0/1

- ☑ Ridge and Lasso ✗
- ☐ Only l1
- ☐ Only l2
- ☑ Both l1 and l2 ✓
- ☐ Neither of the above

Correct answer

- ☑ Both l1 and l2

✗ 5. In order to avoid overfitting , when small set of features are present. Which technique is preferable?

- [ ] Ridge
- [x] Lasso ✗
- [ ] Cross-validation
- [x] Step-wise Regression ✗

Correct answer

- [x] Cross-validation

✗ 6. Can Gradient Descent get stuck in a local minimum when training a Logistic Regression model? 0/1

- ( ) True
- (•) False ✗

Correct answer

- (•) True

Feedback

If we do not have a bowl shaped convex function, Gradient descent can get stuck at local minimum

✓ 7. Which of the following methods do we use to best fit data in Logistic Regression? 1/1

- ○ Least Square Error
- ◉ Maximum Likelihood ✓
- ○ Jaccard Distance
- ○ Both Least Square Error and Maximum Likelihood

✓ 8. Which of the following evaluation metrics cannot be applied in case of logistic regression output to compare with target? 1/1

- ◉ Mean-Squared-Error ✓
- ○ Logloss
- ○ AUC-ROC
- ○ Accuracy

The following table of coefficients is the output of a logistic regression analysis which explores the factors underlying whether or not a student is deemed eligible for free school meals:

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Mother's highest qualification			555.083	5	.000	
hiquamum (Degree or equivalent)	-1.763	.181	95.317	1	.000	.172
hiquamum (HE below degree level)	-1.785	.138	167.001	1	.000	.168
hiquamum (A level or equiv)	-1.746	.120	213.446	1	.000	.174
hiquamum (GCSE grades A-C or equiv)	-1.397	.075	343.134	1	.000	.247
hiquamum (Other)	-1.131	.102	121.958	1	.000	.323
Socio-Economic Class			658.210	2	.000	
Managerial & Pro	-2.120	.105	407.737	1	.000	.120
Intermediate	-1.407	.071	393.136	1	.000	.245
Single Parent	1.645	.063	688.176	1	.000	5.181
Constant	-.389	.051	58.936	1	.000	.678

✗ 9. How many explanatory variables are included in this model? 0/1

- ○ 3
- ○ 5
- ○ 9
- ◉ 10 ✗
- ○ 11

Correct answer

- ◉ 9

Feedback

There are dummy variables for Mother's qualification(5 dummy) and socio economic class(2). Then there are single parent and intercept so 9 in total

✗ 10. Using the above table, how much more likely is that a student from a single parent family will be eligible for free school meals compared to a student not from a single parent family? Please type your answer to the nearest whole number. ···/1

1.65 ✗

Correct answer

5

Feedback

From the last column Exp(B) we see that single parent is 5 times more likely

✗ 10. It is approximately 8 times more likely that someone from the baseline SEC category 'Routine, semi-routine or unemployed' will be eligible for free school meals than someone from the category 'Managerial & professional'. True or false ? 0/1

- ○ True
- ◉ False ✗

Correct answer

- ◉ True

Feedback

The odds ratio, Exp(B), for 'Managerial & professional' homes is 0.12 relative to the reference category of 'semi-routine, routine or LT unemployed' homes. The complement of this OR is 1/.12=8.33.

✗ 11. Please tick all statements which are true 0/1

- [x] Those from the "Routine, semi-routine and unemployed" SEC category are most likely to be eligible for free school meals regardless of maternal education ✓
- [x] Of all those from the 'Routine, semi-routine and unemployed" SEC category , those who have a mother with a degree are least likely to be eligible for free school meals ✗
- [x] There appears to be an interaction between SEC and maternal education in relation to free school meal eligibility ✓
- [x] There does not appear to be an interaction between SEC and maternal education in relation to free school meal eligibility ✗

Correct answer

- [x] Those from the "Routine, semi-routine and unemployed" SEC category are most likely to be eligible for free school meals regardless of maternal education
- [x] There appears to be an interaction between SEC and maternal education in relation to free school meal eligibility

✗ 12. The 5 assumptions for linear regression are (1) independence (2) linearity (3) normality (4) homogeneity of variance & (5) non-multicollinearity. Logistic Regression does not require: 0/1

- ○ 1, 2, 3
- ○ 2, 3, 4
- ◉ 3, 4 , 5 ✗
- ○ Is flexible with all 5
- ○ Is strict with all 5

Correct answer

- ◉ 1, 2, 3

Feedback

Logistic regression assumes
Linearity between independent variables and log odds not dependent and independent variables
No Normality of error terms except for very few cases

✗ 13. The logit function (given as l(x)) is the log of odds function. What could be the range of logit function in the domain x = [0,1]? 0/1

- ○ (− ∞ , ∞)
- ○ (0 , 1)
- ○ (0, ∞)
- ◉ (- ∞, 0) ✗

Correct answer

- ◉ (− ∞ , ∞)

✗ 14. Which of the following option is true? 0/1

- ○ Linear Regression errors have to be normally distributed but not for Logistic Regression
- ○ Logistic Regression errors have to be normally distributed but not for Linear Regression
- ◉ Both Linear Regression and Logistic Regression error values have to be normally distributed ✗
- ○ Both Linear Regression and Logistic Regression errors need not be normally distributed

Correct answer

- ◉ Both Linear Regression and Logistic Regression errors need not be normally distributed

Feedback

It is often considered to be an assumption for Linear Regression, however it is desirable to have normally distributed errors in some cases for easier model comprehension

✓ 15. Adjusted R-squared is always expected to be _____ than R-squared 1/1

- ◉ Lower ✓
- ○ Higher
- ○ Equal
- ○ Can't say

✗ 16. Given the probability of an event occuring is p, what is the corresponding logit function ···/1

p/1-p ✗

Correct answer

p/(1-p)

✗ 17. What is the relation between logit and logistic function? ···/1

logarithmic ✗

Correct answer

Inverse

Feedback

Logit and logistic are inverse functions logit(p) = 1/logistic(p)

✗ 18. How do you improve the validation set score on a logistic regression model? 0/1

- ○ Data resampling
- ◉ Get more training data ✗
- ○ Tune learning rate
- ○ All of the above
- ○ Can't say

Correct answer

- ◉ All of the above

Feedback

All the methods work but they do not guarantee considerable performance gain. Sometimes all of them combined work best, sometimes only 1 of them could work. It is more dependent on the problem at hand and domain expertise

✗ 19. Logistic regression is sensitive to outliers 0/1

- ○ True
- ◉ False ✗

Correct answer

- ◉ True

Feedback

Having outliers would change the parameters in the equation

✗ 20. Does including more features guarantee increase in train performance? 0/1

- (●) True ✗
- ( ) False
- ( ) Can't say

Correct answer

- (●) Can't say

Feedback

It depends on how much information the new feature carries