Summer Analytics 2022

A primer course on Data Science and Machine Learning

Consulting & Analytics Club
IIT Guwahati

Week 3 Quiz (Graded)

Total points **17/30**

Hope that you've gone through the course content for week-3 as well as the covered assignment before attempting the quiz.

- This form accepts the solution only once, so make sure you don't press the submit button accidentally. No requests will be entertained.
- Use the SAME email ID which you used for registering for Summer Analytics 2022.
- Please follow the honor code, which otherwise may lead to harsh actions being taken.
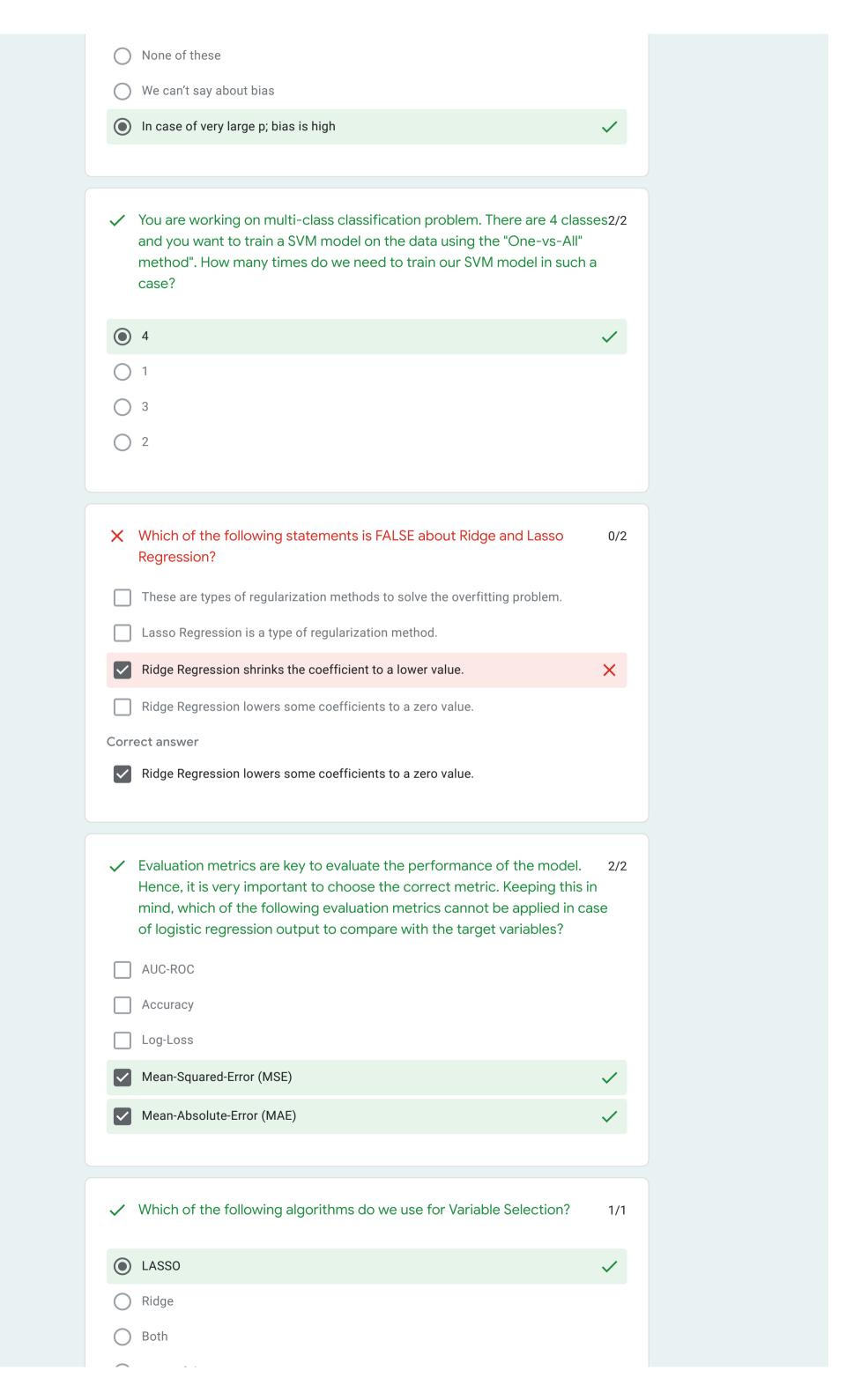
All the best :)

Email *

hridayagrawal0102@gmail.com

0 of 0 points

Name *

Hriday Agrawal

Enrolment ID *

This is a 5-digit number of the form XXXXX. It can be found in the enrolment confirmation mail sent to you.

41395

Are you from IIT Guwahati ? *

○ Yes

◉ No

If you are from IIT Guwahati , provide your roll no.

Graded Quiz   **9 of 21 points**
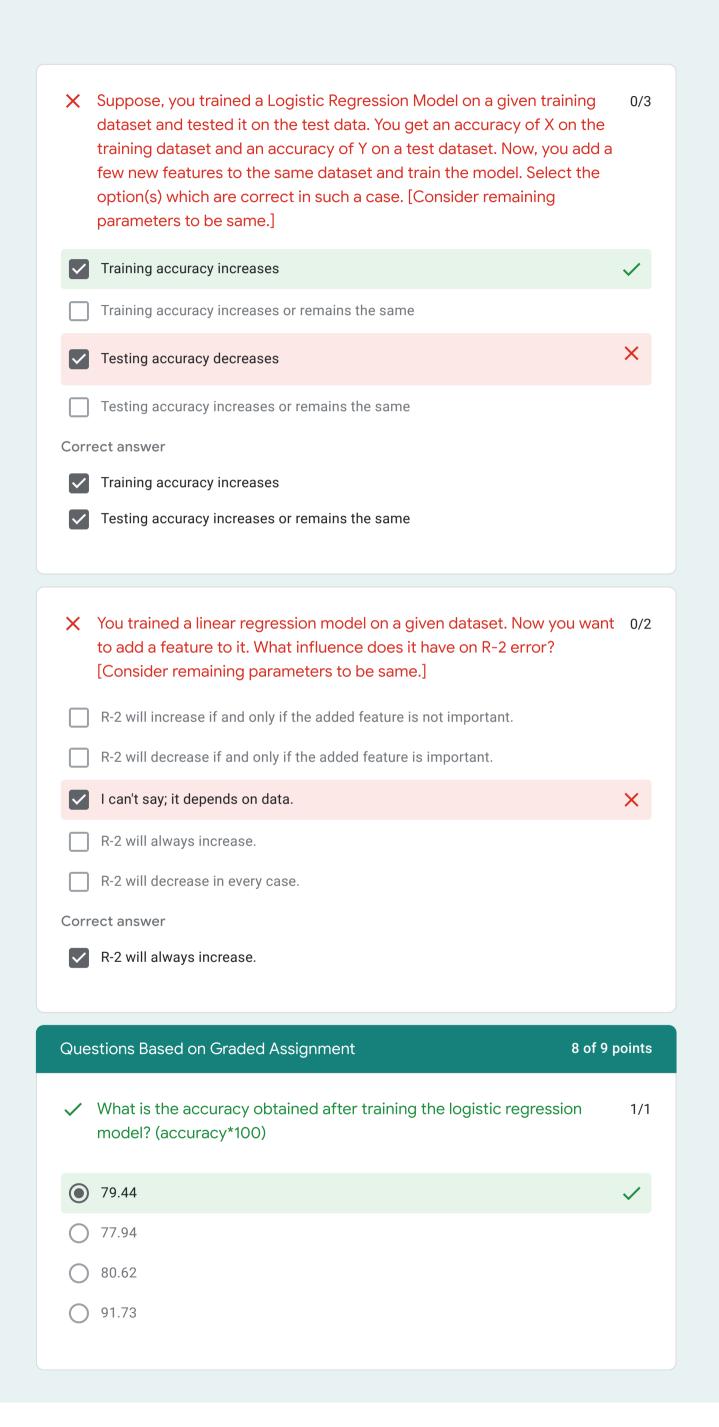
✓ You are given a dataset for which you have to fit a complex regression model. Suppose you are using Ridge Regression with a penalty p. Which of the following options describe bias in the best way?   2/2

○ In case of very large p, bias is low

○ None of these

○ We can't say about bias

⦿ In case of very large p; bias is high ✓

✓ You are working on multi-class classification problem. There are 4 classes **2/2** and you want to train a SVM model on the data using the "One-vs-All" method". How many times do we need to train our SVM model in such a case?

⦿ 4 ✓

○ 1

○ 3

○ 2

✗ Which of the following statements is FALSE about Ridge and Lasso **0/2** Regression?

☐ These are types of regularization methods to solve the overfitting problem.

☐ Lasso Regression is a type of regularization method.

☑ Ridge Regression shrinks the coefficient to a lower value. ✗

☐ Ridge Regression lowers some coefficients to a zero value.

Correct answer

☑ Ridge Regression lowers some coefficients to a zero value.

✓ Evaluation metrics are key to evaluate the performance of the model. **2/2** Hence, it is very important to choose the correct metric. Keeping this in mind, which of the following evaluation metrics cannot be applied in case of logistic regression output to compare with the target variables?

☐ AUC-ROC

☐ Accuracy

☐ Log-Loss

☑ Mean-Squared-Error (MSE) ✓

☑ Mean-Absolute-Error (MAE) ✓

✓ Which of the following algorithms do we use for Variable Selection? **1/1**

⦿ LASSO ✓

○ Ridge

○ Both

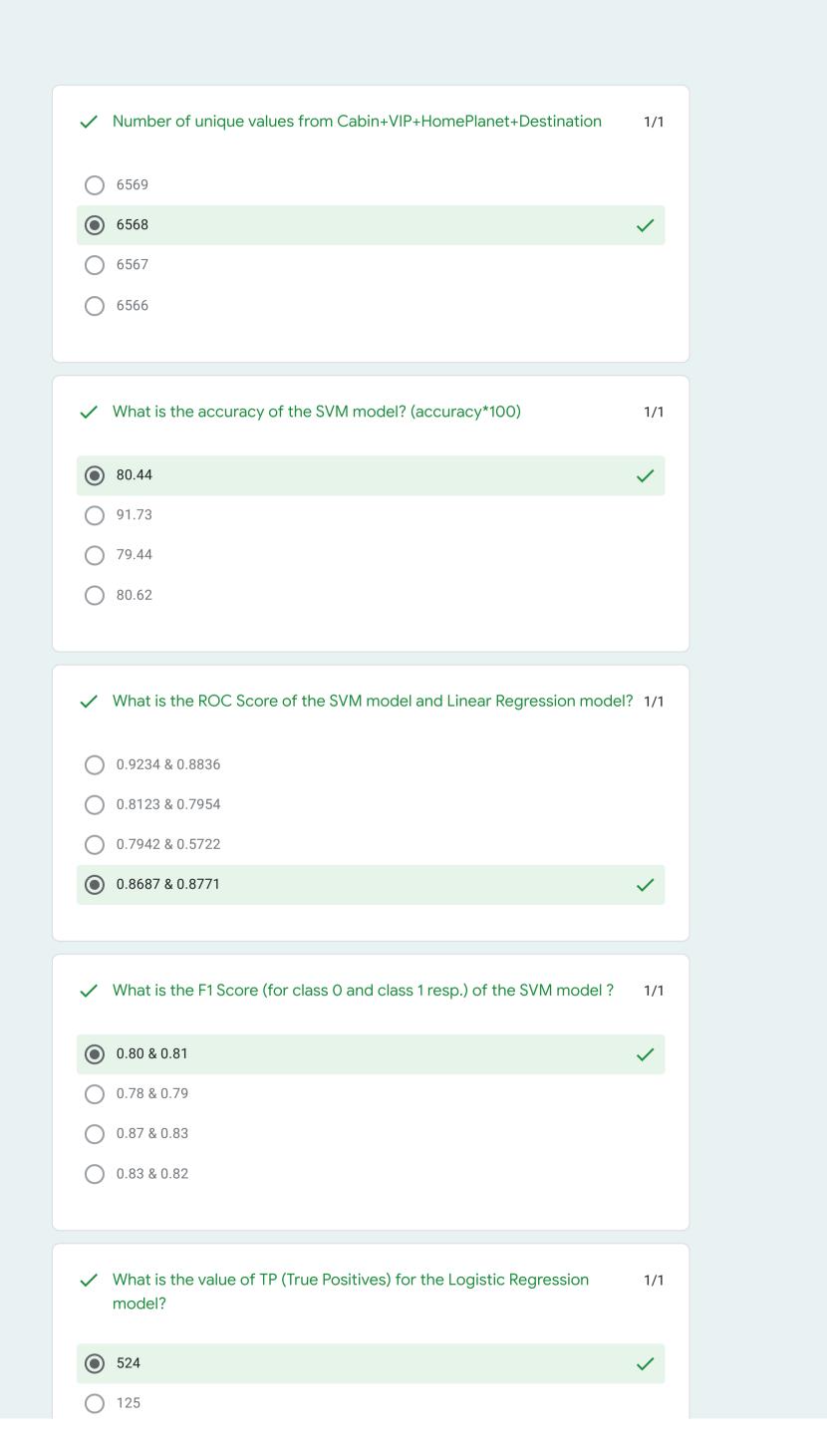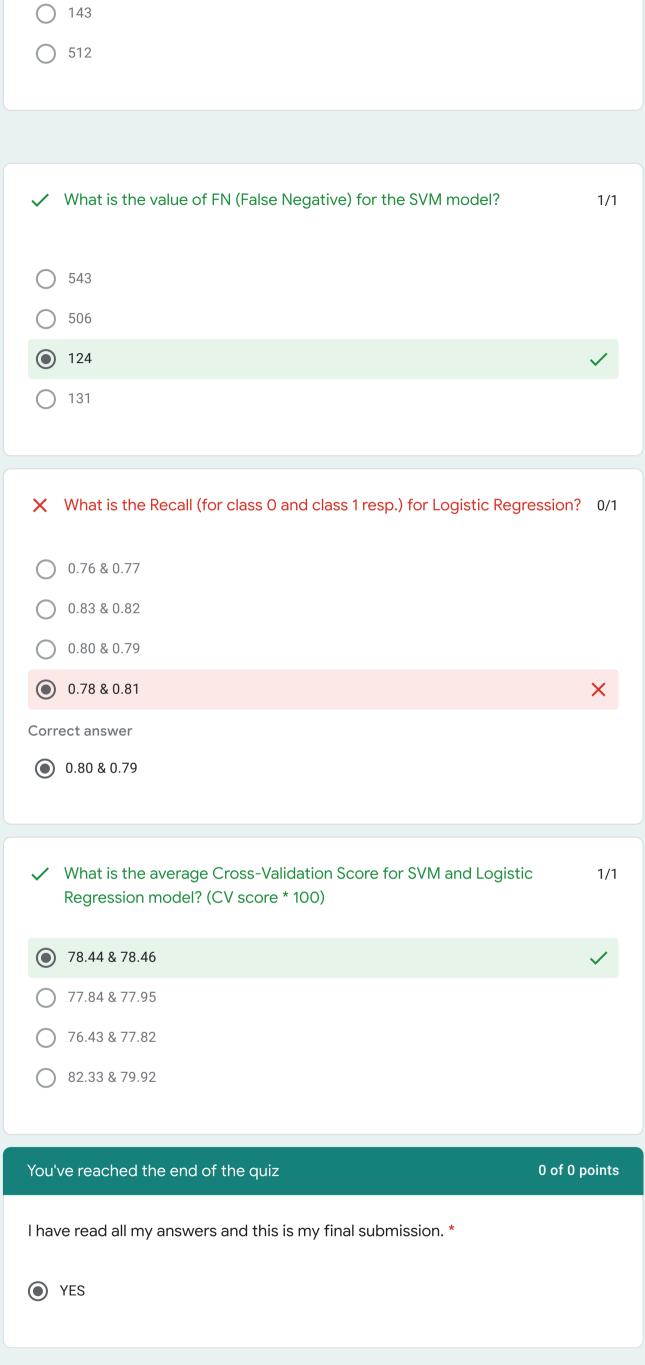○ None of these

✓ Suppose you got a job at YouTube HQ. Your job is to build a classifier so that videos are safe for kids. Your aim should be   2/2

○ High Recall

◉ High Precision   ✓

○ High Accuracy

○ None of the above

✗ You are given a task at hand to train a SVM (Support Vector Machine) with an RBF kernel. You observe that the model is underfitting the training data. What should one change?   0/2

◉ Increase Gamma and Decrease C   ✗

○ Decrease Gamma and Increase C

○ Increase Gamma and Increase C

○ Decrease Gamma and Decrease C

Correct answer

◉ Increase Gamma and Increase C

✗ Feature Normalization is a scaling technique in which values are shifted and rescaled so that they end up ranging between 0 and 1. We usually use it before using a Gaussian Kernel in a SVM model. What is true about feature normalization?   0/3

☐ We do feature normalization so that the new features will dominate the other

☑ Sometimes, it is not feasible to use feature normalization such as in the case of categorical variables.   ✓

☑ Feature normalization always helps when we use Gaussian Kernel in SVM   ✗

☐ All of these

Correct answer

☑ We do feature normalization so that the new features will dominate the other

☑ Sometimes, it is not feasible to use feature normalization such as in the case of categorical variables.

✕ Suppose, you trained a Logistic Regression Model on a given training dataset and tested it on the test data. You get an accuracy of X on the training dataset and an accuracy of Y on a test dataset. Now, you add a few new features to the same dataset and train the model. Select the option(s) which are correct in such a case. [Consider remaining parameters to be same.] 0/3

- [x] Training accuracy increases ✓
- [ ] Training accuracy increases or remains the same
- [x] Testing accuracy decreases ✕
- [ ] Testing accuracy increases or remains the same

Correct answer

- [x] Training accuracy increases
- [x] Testing accuracy increases or remains the same

✕ You trained a linear regression model on a given dataset. Now you want to add a feature to it. What influence does it have on R-2 error? [Consider remaining parameters to be same.] 0/2

- [ ] R-2 will increase if and only if the added feature is not important.
- [ ] R-2 will decrease if and only if the added feature is important.
- [x] I can't say; it depends on data. ✕
- [ ] R-2 will always increase.
- [ ] R-2 will decrease in every case.

Correct answer

- [x] R-2 will always increase.

Questions Based on Graded Assignment — 8 of 9 points

✓ What is the accuracy obtained after training the logistic regression model? (accuracy*100) 1/1

- ◉ 79.44 ✓
- ○ 77.94
- ○ 80.62
- ○ 91.73

✓ Number of unique values from Cabin+VIP+HomePlanet+Destination 1/1

○ 6569

◉ 6568 ✓

○ 6567

○ 6566

✓ What is the accuracy of the SVM model? (accuracy*100) 1/1

◉ 80.44 ✓

○ 91.73

○ 79.44

○ 80.62

✓ What is the ROC Score of the SVM model and Linear Regression model? 1/1

○ 0.9234 & 0.8836

○ 0.8123 & 0.7954

○ 0.7942 & 0.5722

◉ 0.8687 & 0.8771 ✓

✓ What is the F1 Score (for class 0 and class 1 resp.) of the SVM model ? 1/1

◉ 0.80 & 0.81 ✓

○ 0.78 & 0.79

○ 0.87 & 0.83

○ 0.83 & 0.82

✓ What is the value of TP (True Positives) for the Logistic Regression model? 1/1

◉ 524 ✓

○ 125

○ 143

○ 512

✓ **What is the value of FN (False Negative) for the SVM model?** 1/1

○ 543

○ 506

⦿ 124 ✓

○ 131

✕ **What is the Recall (for class 0 and class 1 resp.) for Logistic Regression?** 0/1

○ 0.76 & 0.77

○ 0.83 & 0.82

○ 0.80 & 0.79

⦿ 0.78 & 0.81 ✕

Correct answer

⦿ 0.80 & 0.79

✓ **What is the average Cross-Validation Score for SVM and Logistic Regression model? (CV score * 100)** 1/1

⦿ 78.44 & 78.46 ✓

○ 77.84 & 77.95

○ 76.43 & 77.82

○ 82.33 & 79.92

You've reached the end of the quiz	**0 of 0 points**

I have read all my answers and this is my final submission. *

⦿ YES