Module 3 Quiz

1.Question 1

A supervised learning model has been built to predict whether someone is infected with a new strain of a virus. The probability of any one person having the virus is 1%. Using accuracy as a metric, what would be a good choice for a baseline accuracy score that the new model would want to outperform?

Ans 0.99

Correct

1 / 1 point

2.Question 2

Given the following confusion matrix:

	Predicted Positive	Predicted Negative
Condition Positive	96	4
Condition Negative	8	19

Compute the accuracy to three decimal places.

Ans 0.906

Correct

1 / 1 point

3.Question 3

Given the following confusion matrix:

	Predicted Positive	Predicted Negative
Condition Positive	96	4
Condition Negative	8	19

Compute the precision to three decimal places.

Ans 0.923

4.Question 4

Given the following confusion matrix:

	Predicted Positive	Predicted Negative
Condition Positive	96	4
Condition Negative	8	19

Compute the recall to three decimal places.

Ans 0.960

5.Question 5

Using the fitted model `m` create a precision-recall curve to answer the following question:

For the fitted model `m`, approximately what precision can we expect for a recall of 0.8?

(Use y_test and X_test to compute the precision-recall curve. If you wish to view a plot, you can use `plt.show()`)

```
#print(m)

pre,rec,_ = precision_recall_curve(y_test,m.predict(X_test))

plt.plot(rec,pre)

plt.xlabel('Recall')

plt.ylabel('Precision')

plt.ylim([0.0, 1.05])

plt.xlim([0.0, 1.00])
```

```
plt.show()
```
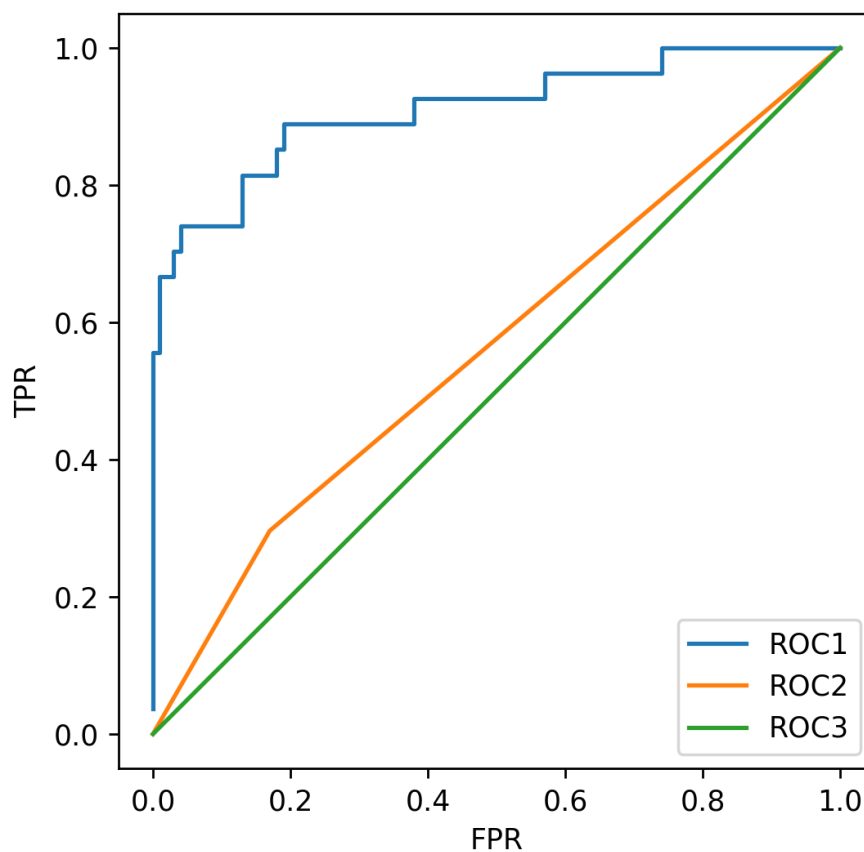
```
RunReset
```

Ans 0.6

Correct

1 / 1 point

6.Question 6

Given the following models and AUC scores, match each model to its corresponding ROC curve.

- Model 1 test set AUC score: 0.91

- Model 2 test set AUC score: 0.50

- Model 3 test set AUC score: 0.56
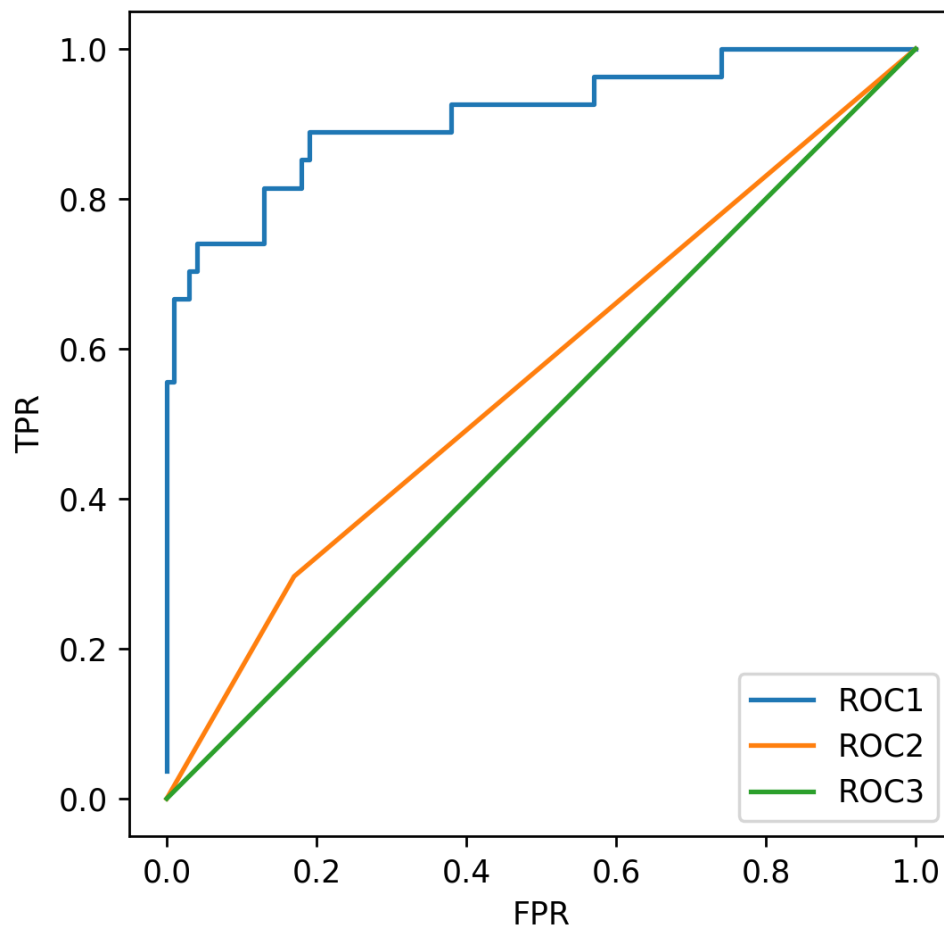

Ans   Model 1: ROC 1
Model 2: ROC 3
Model 3: ROC 2

Correct

1 / 1 point

7.Question 7

Given the following models and accuracy scores, match each model to its corresponding ROC curve.

- Model 1 test set accuracy: 0.91
- Model 2 test set accuracy: 0.79
- Model 3 test set accuracy: 0.72

Ans Not enough information is given.

Correct

1 / 1 point

8.Question 8

Using the fitted model `m` what is the micro precision score?

(Use y_test and X_test to compute the precision score.)

1
2

```
#print(m)
print(precision_score(y_test,m.predict(X_test),average = 'micro'))
```
```
RunReset
0.744

0.744

None
```

Ans 0.744

Correct

`1 / 1 point`

9.Question 9

Which of the following is true of the R-Squared metric? (Select all that apply)

Ans   The best possible score is 1.0

       A model that always predicts the mean of y would get a score of 0.0

Correct

`1 / 1 point`

10.Question 10

In a future society, a machine is used to predict a crime before it occurs. If you were responsible for tuning this machine, what evaluation metric would you want to maximize to ensure no innocent people (people not about to commit a crime) are imprisoned (where crime is the positive label)?

Ans Precision

Correct

`1 / 1 point`

11.Question 11

Consider the machine from the previous question. If you were responsible for tuning this machine, what evaluation metric would you want to maximize to ensure all criminals (people about to commit a crime) are imprisoned (where crime is the positive label)?

Ans Recall

Correct

`1 / 1 point`

12.Question 12

A classifier is trained on an imbalanced multiclass dataset. After looking at the model's precision scores, you find that the micro averaging is much smaller than the macro averaging score. Which of the following is most likely happening?

Ans The model is probably misclassifying the frequent labels more than infrequent labels.

Correct

1 / 1 point

13.Question 13

Using the already defined RBF SVC model `m`, run a grid search on the parameters C and gamma, for values [0.01, 0.1, 1, 10]. The grid search should find the model that best optimizes for recall. How much better is the recall of this model than the precision? (Compute recall - precision to 3 decimal places)

(Use y_test and X_test to compute precision and recall.)

```
#print(m)

parameters = {'gamma' : [0.01, 0.1, 1, 10], 'C' : [0.01, 0.1, 1, 10]}

clf = GridSearchCV(m, parameters, scoring = 'recall')

clf.fit(X_train, y_train)

y_pred = clf.best_estimator_.predict(X_test)

rec = recall_score(y_test, y_pred, average = 'binary')

pre = precision_score(y_test, y_pred, average = 'binary')

print(rec - pre)
```

RunReset

```
0.52

0.52

None
```

Ans 0.52

Correct

1 / 1 point

14.Question 14

Using the already defined RBF SVC model `m`, run a grid search on the parameters C and gamma, for values [0.01, 0.1, 1, 10]. The grid search should find the model that best optimizes for precision. How much better is the precision of this model than the recall? (Compute precision - recall to 3 decimal places)

(Use y_test and X_test to compute precision and recall.)

```
#print(m)

parameters = {'gamma' : [0.01, 0.1, 1, 10], 'C' : [0.01, 0.1, 1, 10]}

clf = GridSearchCV(m, parameters, scoring = 'precision')

clf.fit(X_train, y_train)

y_pred = clf.best_estimator_.predict(X_test)

rec = recall_score(y_test, y_pred, average = 'binary')

pre = precision_score(y_test, y_pred, average = 'binary')

print(pre - rec)
```

RunReset

```
0.15

0.15

None
```

Ans 0.15

Correct