Name: _____
Score: _____ / _____

Quiz #3

Part 1

Which of the following is/are true about boosting trees?

1. It is the method for improving the performance by aggregating the results of weak learners

2. In boosting trees, individual weak learners are independent of each other

A.
1

B.
2

C.
1 and 2

D.
None of the above

Answer Point Value: 1.0 points
Answer Key: A

How do we estimate a bagging output?

A.
Averaging for regression, voting for classification

B.
Regressing for classification, Classifying for regression

C.
Minimizing the loss for regression, minimizing the likelihood for classification

D.
Voting for regression, averaging for classification

Answer Point Value: 1.0 points
Answer Key: A

Which of the following is a reasonable way to select the number of principal components k?
(Recall that n is the dimensionality of the input data and m is the number of input examples.

A.
Choose k to be the largest value so that at least 99% of the variance is retained

B.
Choose k to be 99% of n (i.e., $k = 0.99 * n$, rounded to the nearest integer).

C.
Choose k to be the smallest value so that at least 1% of the variance is retained.

D.
Choose k to be the smallest value so that at least 99% of the variance is retained.

Answer Point Value: 1.0 points
Answer Key: D

Which of the following are recommended applications of PCA? **Select two that apply.**

A.
Preventing overfitting

B.
Data visualization: Reduce data to 2D (or 3D) so that it can be plotted.

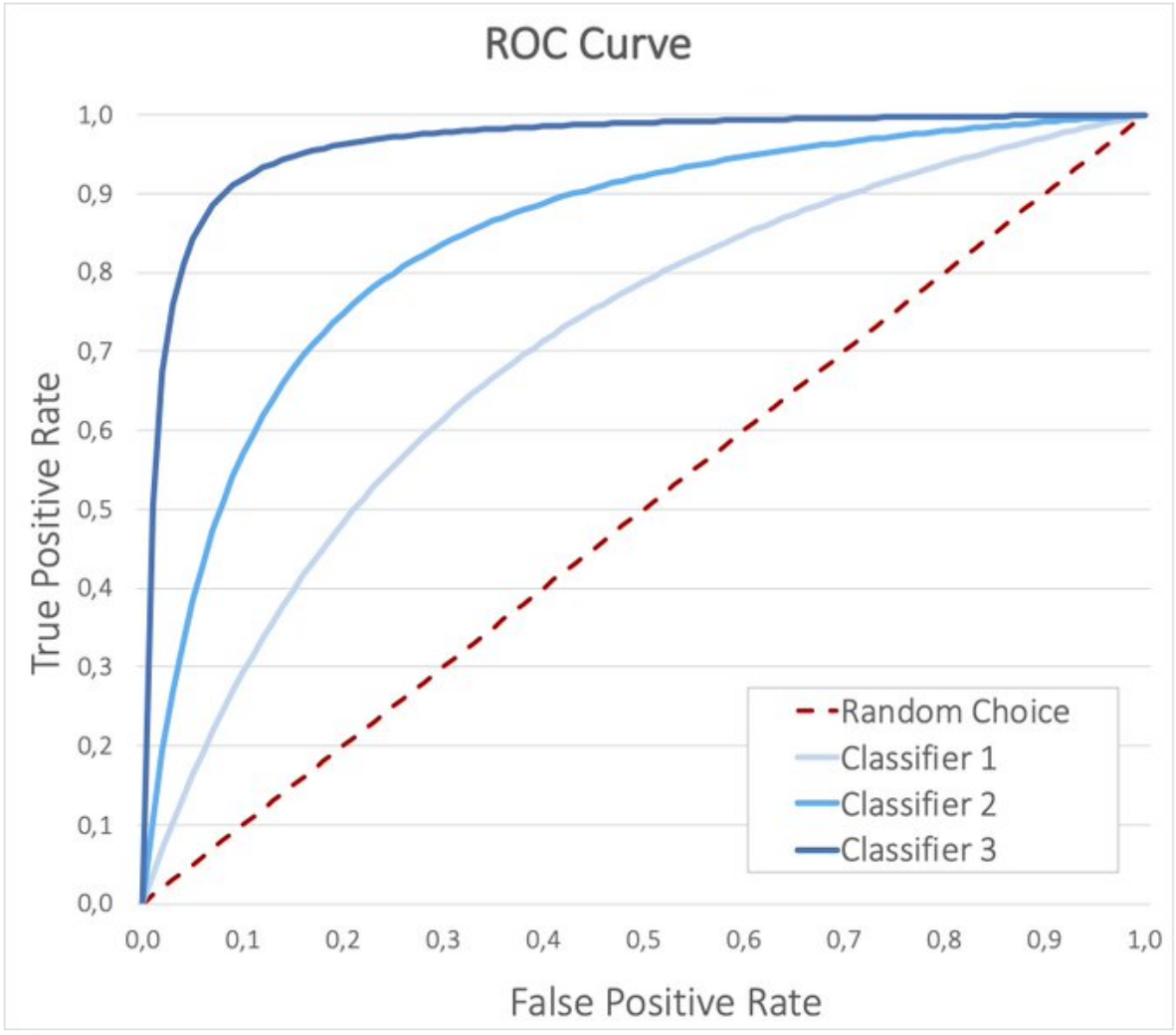
C.
To get more features to feed into a learning algorithm.

D.
Data compression: Reduce the dimension of your data, so that it takes up less memory / disk space.

Answer Point Value: 1.0 points
Answer Key: B,D

Which of the following algorithms would you consider in your final model building on the basis of performance? Suppose you have given the following graph which shows the ROC curve for three different classification models: Classifiers 1, 2, and 3.

Attachments



A. Classifier 1

B. Classifier 2

C. Classifier 3

D. Random Choice

Answer Point Value: 1.0 points
Answer Key: C

In PCA, what does the variance of a principal component indicate?

A.
The degree to which the principal component can be used for classification.

B.
The degree to which the principal component can be used for regression.

C.
The degree of correlation between the principal component and the original data.

D.
The amount of information contained in the principal component.

Answer Point Value: 1.0 points
Answer Key: D

Which of the following algorithm is based on the bagging strategy?

A.
Decision Tree

B.
Random Forest

C.
XGboost

D.
AdaBoost

Answer Point Value: 1.0 points
Answer Key: B

Why would we use a random forest instead of a decision tree?

A.
For lower training error.

B.
To reduce the variance of the model.

C.
For a model that is easier for a human to interpret.

D.
All of the above.

Answer Point Value: 1.0 points
Answer Key: B

In Random forest you can generate hundreds of trees (say T1, T2 …..Tn) and then aggregate the results of these tree. Which of the following is true about individual(Tk) tree in Random Forest?

1. Individual tree is built on a subset of the features
2. Individual tree is built on all the features
3. Individual tree is built on a subset of observations
4. Individual tree is built on full set of observations

A.
1 and 4

B.
2 and 3

C.
2 and 4

D.
1 and 3

Answer Point Value: 1.0 points
Answer Key: D

Which of the following is/are true about Random Forest and AdaBoost methods?

1. Both methods can be used for classification task
2. Random Forest is used for classification whereas AdaBoost is used for regression task
3. Random Forest is used for regression whereas AdaBoost is used for Classification task
4. Both methods can be used for regression task

A.
1

B.
2

C.
3

D.
4

Answer Point Value: 1.0 points
Answer Key: A,D