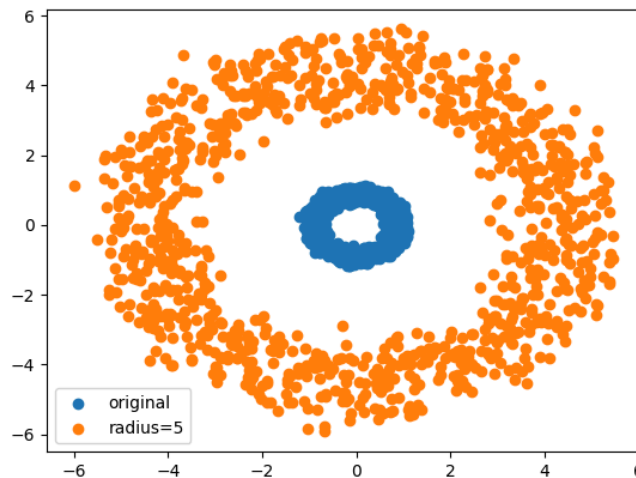


Name:

NetID:

Note: This quiz will be a set of practice problems for the upcoming exam. Complete this at your leisure outside of class but please bring in an **INDIVIDUAL** copy of the quiz **ON PAPER** to turn in at the start of the exam. While doing the work feel free to work with your classmates, but I'd like you all to turn in a copy individually. You don't necessarily need to print and fill out this exact document, just writing the answers on your own piece of paper will be fine.

1.) Which of the three SVM kernels we've seen so far would be best for separating the data points shown in the plot below?



A.) RBF

B.) Linear

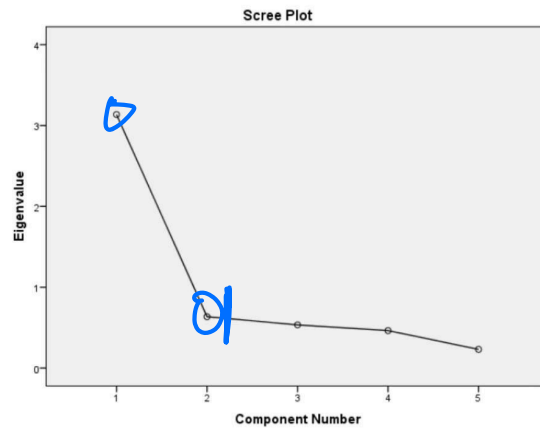
C.) Polynomial

D.) Lagrangian

2.) Is a regression task supervised or unsupervised learning? Explain your reasoning.

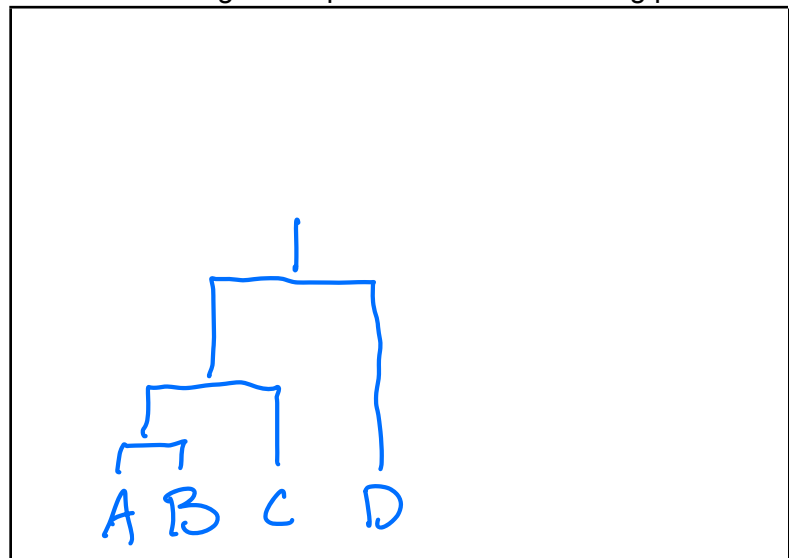
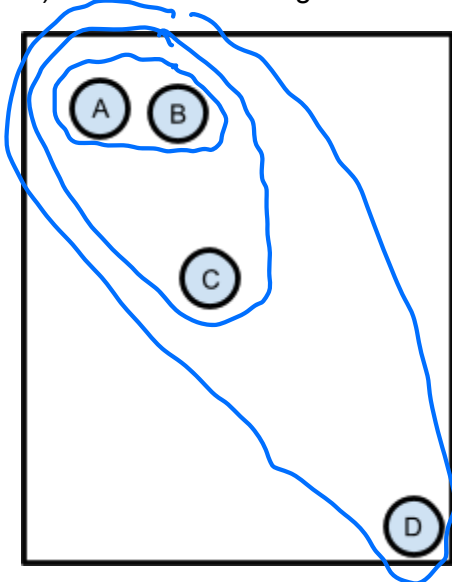
Regression is supervised because we have ground truth labels

3.) When we use PCA, we often get more principal components back than we'd realistically use. Given the scree plot below, indicate how many principal components you'd choose to use. Why don't we use every component?



I'd use the first two principle components because that's where the elbow is. Honestly just one could be fine given that scree plot. PCAs capture less and less variance beyond the first few

4.) Create the dendrogram that hierarchical clustering would produce for the following plot.



5.) What's an example of a task where a ROC curve is used? When would we prefer AUC?

They are good for binary classification tasks  
We'd prefer AUC if we wanted a single number to represent our performance

6.) Explain the difference between Precision, Recall, and the F1 score. Why might we prefer one to any of the others?

Precision - focuses on accuracy of positive predictions, good when false positives are costly  
Recall - focuses on all positive cases, useful when false negatives are costly  
F1 - balance between precision and recall, used when both are important and for unbalanced datasets

7.) Do you prefer the newer lecture style with the ipad or the older chalkboard based one?

