Name HW4

Description

Rubric Detail

Criteria	Levels of Achievement		
	Novice	**Competent**	**Proficient**
Q1-Agglomerative Clustering (Single) **Weight** **5.00%**	**0 %** Not Done	**75 %** Fundamental problem in implementation.	**100 %** Correctly performing the clustering task. Most students will probably use hclust as was advised by the problem.
Q1-Agglomerative Clustering (Average) **Weight** **5.00%**	**0 %** Not Done	**75 %** Fundamental problem in implementation.	**100 %** Correctly performing the clustering task. Most students will probably use hclust as was advised by the problem.
Q1-Agglomerative Clustering (Complete) **Weight** **5.00%**	**0 %** Not Done	**75 %** Fundamental problem in implementation.	**100 %** Correctly performing the clustering task. Most students will probably use hclust as was advised by the problem.
Q1-KMeans-Choice of k **Weight** **4.00%**	**0 %** Not done, or completely out of range.	**80 %** K=2,9,10	**100 %** 3<=k<=8
Q1-KMeans-Reasoning for k **Weight** **8.00%**	**0 %** Not Given.	**80 %** Partially correct reasons	**100 %** For k=3 argument could be historical/economical. Include some mathematical reasoning if arguing k above 3. For instance, the mathematical reasoning could be the elbow plot or any model selection test.

Criteria	Levels of Achievement		
	Novice	Competent	Proficient
Q1-KMeans-Clustering **Weight** **3.00%**	**0 %** Not Given.	**80 %**	**100 %** See if results seem reasonable.While K can vary, similar countries are supposed to be in same group.
Q2-Splitting data **Weight** **10.00%**	**0 %** Not Done	**80 %** The d value was not reported, or an unfit method of splitting resulted in the model being recognizably mistrained.	**100 %** Report method for splitting; d value (the chunks size) and amount of overlap. Expected range 10<=d<=32. d can be most values as long as accuracy is okay (>=60%)
Q2-Making clustering using pieces from training data **Weight** **10.00%**	**0 %** Not Done.	**90 %** Picking unreasonable k values.	**100 %** Reasonable k (50<=k<=600) should be picked for this stage using K means or hierarchical k means. The student is supposed to figure out that the model needs a relatively large k to perform well, and pick a neither too large nor too small value..
Q2-Turning each training sample into histogram based on train clustering **Weight** **10.00%**	**0 %** Not Done.	**80 %** Incorrect translation into histograms.	**100 %** For each training signal sample: 1)Divide the signal into chunks as before 2)Find the closest Centroid to each chunk 3)Represent the signal by a series of closest centroid indices 4) Turn the series of indices into a histogram.
Q2-Building Classifier using training histograms **Weight** **10.00%**	**0 %** Not Done.	**80 %** Building a classifier and incorrectly training it.	**100 %** Any kind of classifier is acceptable as long as implemented correctly. Most students will likely use random forests since the problem advised so, but any other classifiers are also acceptable.

	Levels of Achievement		
Criteria	**Novice**	**Competent**	**Proficient**
Q2-Turning testing samples into histograms using train clustering, and classifying the historgrams and evaluating the test accuracy **Weight** **20.00%**	**0 %** Not Done.	**90 %** Partially correct testing procedure.	**100 %** The test samples should be turned into histograms using the trained centroids.
Q2-Trying other parameters and reporting best accuracy **Weight** **10.00%**	**0 %** Not Done.	**80 %** Partial correctness.	**100 %** Trying at least one other parameter setting. And getting a reasonable accuracy (>=60%).

View Associated Items

Print Close Window