Quiz Feb.04

273A
• Bob and Alice are playing the “data compression game” to figure out how complex of a model they can afford for a particular dataset of N items.

• Bob proposes a model with K real valued parameters. He argues that because the parameters are real valued he needs to quantize the parameter values into L bins.

• He now argues: for fixed dataset size N, if I use more bins (larger L) then I........:

A) Can also use more parameters (larger K) for the optimal compression.

B) Can use less parameters (smaller K) for the optimal compression.

C) L has no influence on K, I can simply take arbitrarily small bins.
• K-means is guaranteed to converge in a finite number of iterations. (This means, there will be no more change in parameters or assignments)

A) True
B) False

• K-means is guaranteed to converge the global minimum of the cost function.

A) True
B) False
• Bagging is not expected to work well when:

A) The individual models are too simple.

B) The individual models are too complex

C) It always works when you use bootstrap samples.
• Imagine you are learning a decision tree from training data and on the current branch you have used all the (discrete) attributes but you end up with training samples of both classes in the same leaf node. Then:

A) This is not possible. You have a bug in your code.

B) You need to pick an attribute that is different from the last one and apply it again.

C) You give up: test cases that end up in that bucket will not be classified

D) You make a decision on test examples based on the parent node