

# CS7641 ML Problem Set II by Georgia Tech

*T. Ruzmetov*

*December 5, 2017*

*Ownership of the following report developed as a result of assigned institutional effort, an assignment of the CS 7641 Machine Learning course shall reside with GT and the instructors of this class. If the document is released into the public domain it will violate the GT Honor Code*

## Problem 1

You have to communicate a signal in a language that has 3 symbols A, B and C. The probability of observing A is 50% while that of observing B and C is 25% each. Design an appropriate encoding for this language. What is the entropy of this signal in bits?

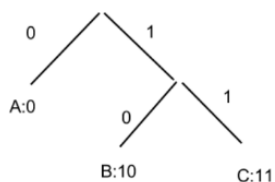


Figure 1:

a.

b.  $S = -\sum_i P_i \log_2 P_i = -(0.5 \log_2 0.5 + 0.25 \log_2 0.25 + 0.25 \log_2 0.25) = 1.5$

## **Problem 2**

Show that the Kmeans procedure can be viewed as a special case of the EM algorithm applied to an appropriate mixture of Gaussian densities model.

3. Plot the direction of the first and second PCA components in the figures given.

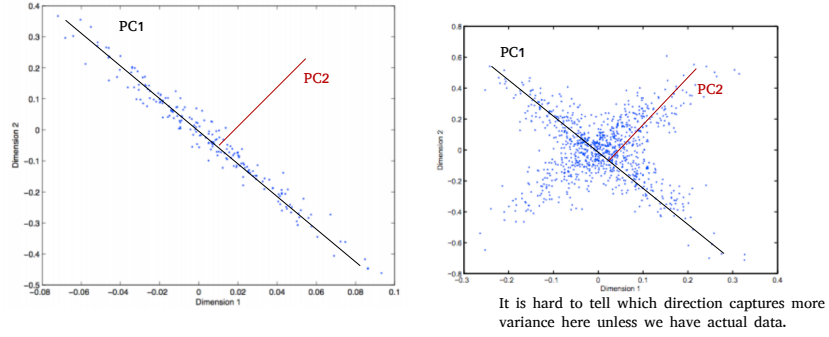


Figure 2: Here we show 6x6(left plot) easy and 13x13(right plot) hard grid world problem topologies. As depicted, grey circle is the agent and blue square at top right corner is final destination.

### Problem 3

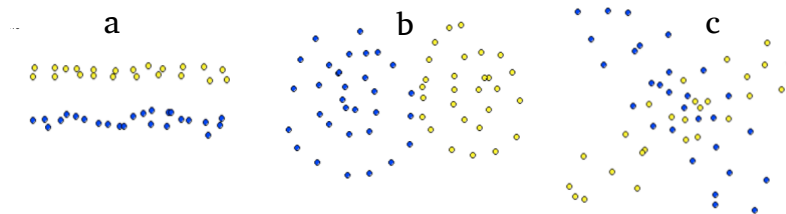


Figure 3:

#### Problem 4

Which clustering method(s) is most likely to produce the following results at  $k = 2$ ? Choose the most likely method(s) and briefly explain why it/they will work better where others will not in at most 3 sentences. Here are the five clustering methods you can choose from:

- Hierarchical clustering with single link
- Hierarchical clustering with complete link
- Hierarchical clustering with average link
- K-means
- EM

## Problem 7

Consider the following simple grid world problem. (Actions are N, S, E, W and are deterministic.)

Our goal is to maximize the following reward:

1. 10 for the transition from state 6 to G
2. 10 for the transition from state 8 to G
3. 0 for all other transitions

S	2	3
4	5	6
7	8	G