CSCI 59000 Machine Learning
Assignment 4
Due: 4/27/2017

1. (25 points) Circle True or False for the following statements. **Write one sentence justification for your answer.**

(True/False) For a SVM classifier with C>0 a sample with the corresponding KKT multiplier equal to 0 indicates an unbounded support vector

(True/False) For a SVM classifier with C>0, $yf(x)<1$ implies x is a bounded support vector

(True/False) A polynomial kernel of degree p has a larger VC dimension than a linear kernel.

(True/False) When two classes are Normally-distributed with equal covariance matrices the resulting Bayes classifier can be expressed as a discriminant function.

(True/False) Reducing the bias of an estimator also reduces its variance.

(True/False) If two random variables are not correlated then they are independent.

(True/False) ln(x) is a convex function.

(True/False) In the EM algorithm the lower bound of the likelihood function is maximized at each iteration.

(True/False) Least square regression and Fisher's linear discriminant both use quadratic loss function.

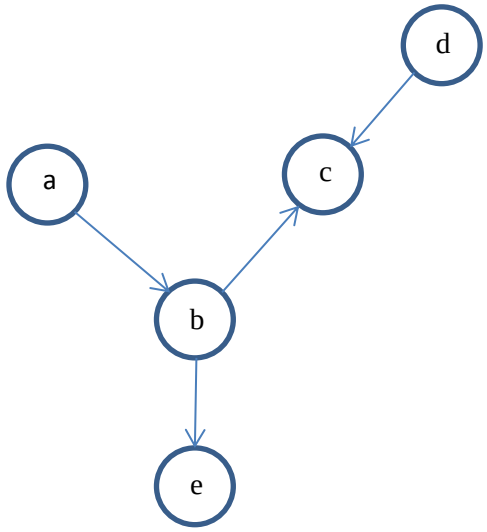
(True/False) The empty set is the smallest possible sigma field.

2. (25 points) Let class 1 and class 2 data be distributed according to mixture distributions $N(-2,2)$ and $N(2,2)$, respectively. Find the posterior probability p(class 1|x=0), i.e. the probability that x=0 belongs to class 1, when the prior probabilities of class 1 and class 2 are as follows.

a.) P1=0.5, P2=0.5
b.) P1=0.8, P2=0.2
c.) P1=0.2, P2=0.8

3. (25 points). Refer to the directed graph below for parts a through f.

a) (4 points) Write p(a,b,c,d,e) in terms of the conditional and prior probabilities of each node.
b) (2 points) (True,False) a is independent of c
c) (2 points) (True,False) a is independent of d
d) (2 points) (True,False) c is independent of e
e) (5 points) Convert this graph to a factor graph.
f) (8 points) Describe a message passing algorithm to compute p(b).
g) (2 points) (Tue, False) Is message passing possible for graphs with loops? Justify your answer with a sentence or two.



4. Given the vocabulary={the, machine, learning, qualifier, exam, I, will, pass, fail, is, easy, hard} and the following topics defined over this vocabulary:

Topic1 $\sim \text{Mult}(\mu_1)$
Topic2 $\sim \text{Mult}(\mu_2)$
Topic3 $\sim \text{Mult}(\mu_3)$
Topic4 $\sim \text{Mult}(\mu_4)$

where μ_i's are as follows

	the	machine	learning	qualifier	exam	I	will	pass	fail	is	easy	hard
μ_1	0.2	0.2	0.2	0.2	0.2	0	0	0	0	0	0	0
μ_2	0	0	0	0	0	0.25	0.25	0	0	0.5	0	0
μ_3	0	0	0	0	0	0	0	0.7		0	0.3	0
μ_4	0	0	0	0	0	0	0	0	0.3	0	0	0.7

We generate a phrase with 7 words using the topic proportion vector β=[0.5 0.2 0.2 0.1].

Rank the following phrases from most likely (1) to least likely (5) to be generated by this generative model. Justify your rationale for ranking the phrases.

Phrase 1: the machine learning qualifier exam is easy

Phrase 2: the machine learning qualifier exam is hard

Phrase 3: I will pass machine learning qualifier exam

Phrase 4: I will fail machine learning qualifier exam

Phrase 5: the exam is hard I will pass