Name: ____

Problem	Points	Score
1(a)	20	
1(b)	10	
2(a)	10	
2(b)	20	
2(c)	10	
3(a)	10	
3(b)	10	
3(c)	10	
Total	100	

Notes:

(1) The exam is closed books and notes.

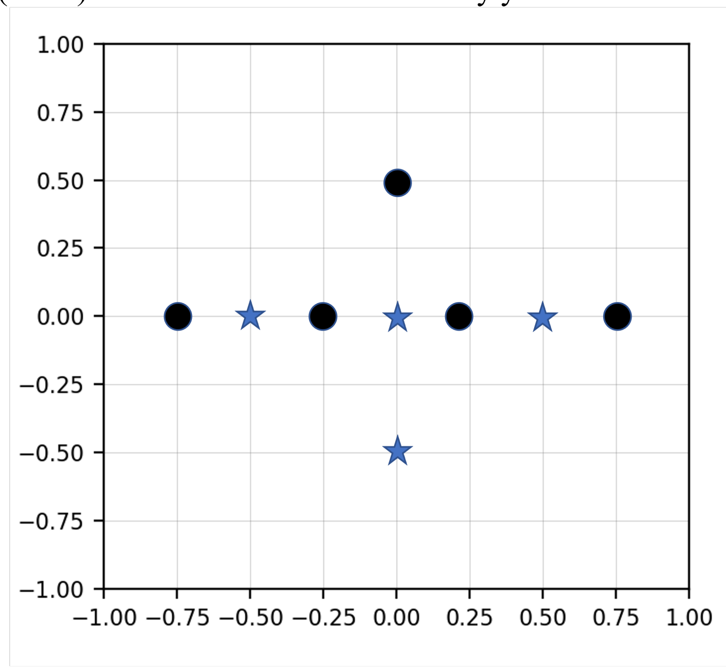
(2) Please clearly indicate your answer to the problem.

(3) Note that ungrammatical sentences, incoherent statements, or general illegible scratches will get zero credit.

(4) If I can't read or follow your solution, it is wrong, and no partial credit will be awarded.

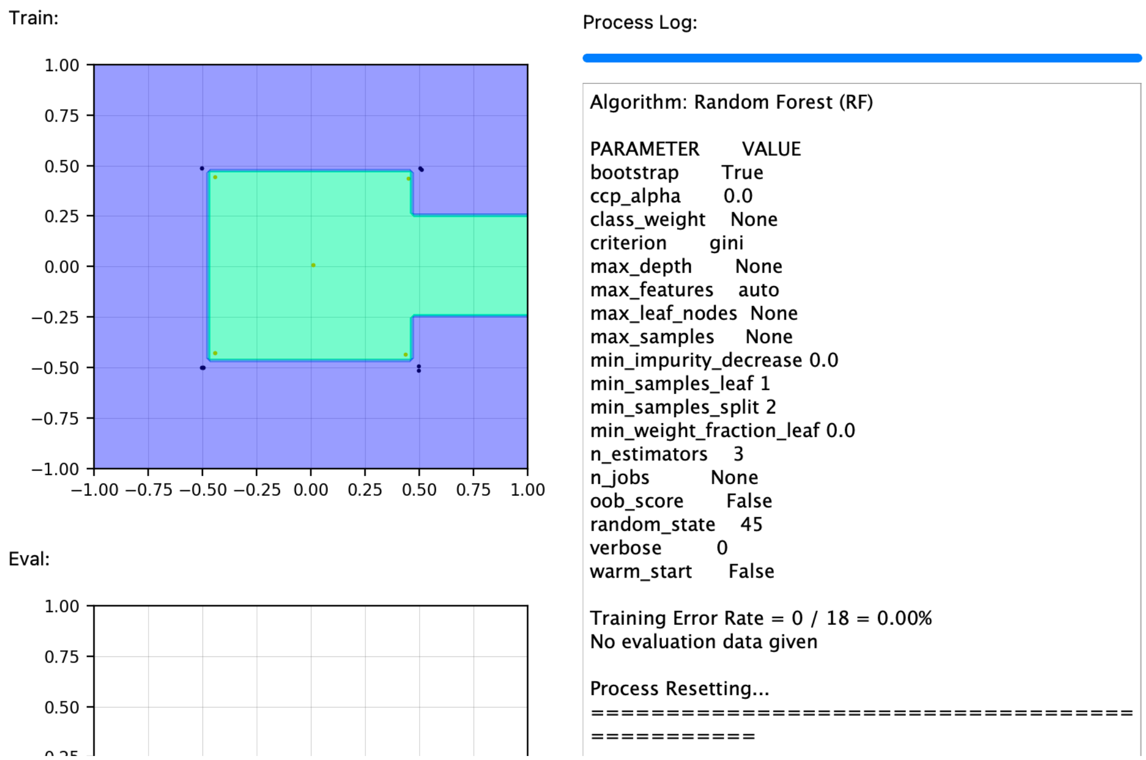
(40 pts) Problem No. 1: Given the data shown below:

(a) (20 pts) Draw the decision surface you would obtain if you applied the k-nearest neighbor algorithm (KNN) with $K = 3$ to this data. Justify your result with a detailed explanation.



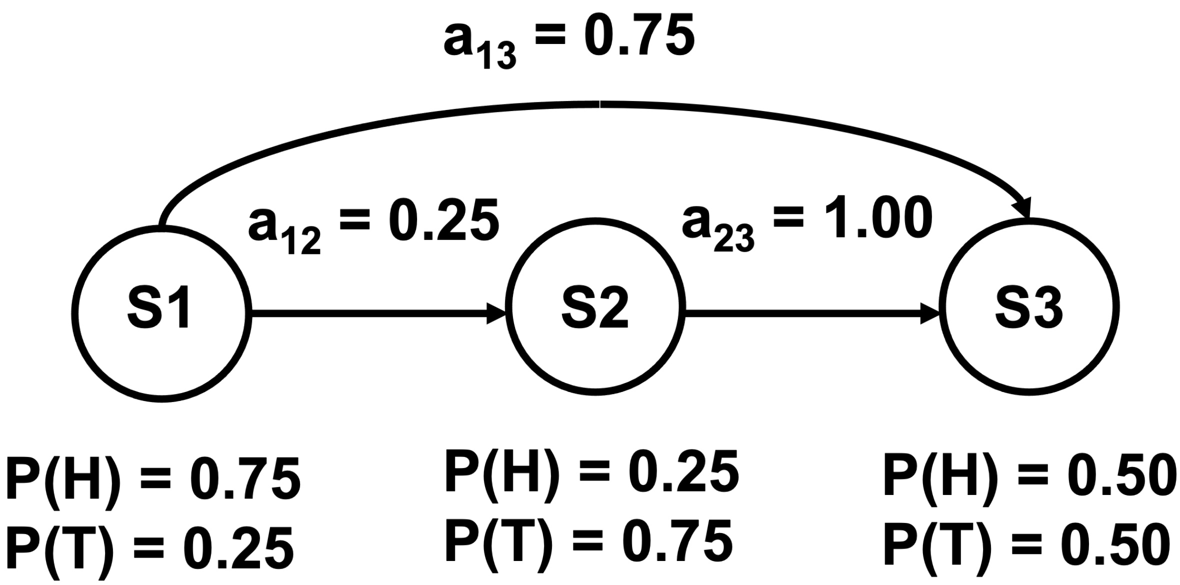
Justification:

(b) (10 pts) I was playing around with IMLD before the exam and discovered a bug when comparing results with JMP. Consider the result shown below:



This was generated with Random Forests using 3 trees. Do you agree with IMLD? Justify your answer.

Problem No. 2: Given the hidden Markov model shown below:



You must start in state S1 and end in state S3. Hence, the initial state probabilities for S_1, S_2 and S_3 are "1.0, 0.0, 0.0" respectively.

(a) (10 pts) Is this actually a "hidden" model. Explain.

(b) (20 pts) Assume you are given the training sequences: “HH”, “HT”, “TH”, “TT”, “HHH”, “TTT”, “HTHT”, and “THTH”. Reestimate the transition probabilities a_{12} and a_{13}.

(c) (10 pts) If you were to use this model to randomly generate data, what is the average duration of the sequences produced?

Problem No. 3: A discrete random variable, X, has a probability mass function (pmf):

$$p_k = \begin{Bmatrix} 1/3 & 0 \le k < 2 \\ 2/3 & 2 \le k < 4 \end{Bmatrix}.$$

A similar random variable, Y, has a probability mass function:

$$p_k = \begin{Bmatrix} 1/4 & 0 \le k < 2 \\ 1/4 & 2 \le k < 4 \end{Bmatrix}.$$

Equations you might find useful for this problem include:

$$H(X) = -\sum_{x \in X} p(x)\, log\, p(x)$$

$$H(X,Y) = -\sum_{x \in X} \sum_{y \in Y} p(x,y)\, log\, p(x,y)$$

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y)\, log \frac{p(x,y)}{p(x)p(y)} = \sum_{x \in X} \sum_{y \in Y} p(x,y)\, log \frac{p(x|y)}{p(x)}$$

$$I(X;Y) = H(X) - H(X|Y)$$

$$I(X;Y) = H(Y) - H(Y|X)$$

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

(a) (10 pts) Compute the entropy of X and Y. Explain why your answers makes sense.

(b) (10 pts) Assume the joint distribution between X and Y is a uniform distribution: $p(x, y) = 1/16$. Compute the mutual information. Justify your answer.

(c) (10 pts) Suggest a shape for the joint distribution that would increase the mutual information. Justify your answer.