Name: _____

Problem	Points	Score
1(a)	30	
1(b)	10	
1(c)	10	
2(a)	15	
2(b)	15	
2(c)	10	
2(d)	10	
Total	100	

Notes:

(1) The exam is closed books and notes except for one double-sided sheet of notes.

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

(3) Note that ungrammatical sentences, incoherent statements, or general illegible scratches will get zero credit. Please carefully explain your solutions in well-written English.

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

(50 pts) Problem No. 1: Consider a three-state discrete HMM model where each state can output one of two symbols, H or T, with equal probability. The first state is a start state, meaning all sequences must pass through this state. The last state is a stop state, meaning all sequences must terminate on this state. State no.1 is connected to states nos. 2 and 3. State 2 is connected to itself and state no. 3. State no. 3 is a terminal state and has no other connections. Assume uniform distributions for the transition probabilities (e.g., 0.5 for each of the two transition probabilities leaving state no. 1; 0.5 for each transition from state no. 2).

Assume an output symbol is emitted in state 1. When you transition into state 2, a second symbol is emitted. If you transition back to state 2, a second symbol is emitted. State 3 does not emit a symbol. Therefore, a transition from state 2 to state 3 does not produce an output. Therefore, the model can produce observation sequences as short as one symbol, and infinitely long as well.

(a) What is the average duration of a sequence output from this model? How would you describe the shape of this distribution?

(b) What is the probability this model produced the following sequences: "H", "HH", "HHH", "HHHH"?

(c) Train the model using the following data: "H", "T", "HH", "TT", "HHH", "TTT". Is this a hidden Markov model? Explain.

(50 pts) Problem No. 2: You are given two training data sets:

Class 1: [0.0, 0.0], [1.0, 0.0], [1.0, 1.0], [0.0, 1.0];

Class 2: [0.25, 0.25], [-1.0, 0.0], [-1.0, -1.0], [0, -1.0].

Your blind evaluation set consists of "Class 1: [0.5,0.5]" and "Class 2: [-0.5, -0.5]".

(a) Design a decision tree to classify this data. What is the probability of error?

(b) Design a kNN algorithm to classify this data using k=2 (walk through the steps of classifying each point using the training data). What is the probability of error?

(c) Design a Support Vector Machine to classify this data (keep it simple). What is the probability of error?

(d) Compare and contrast these approaches. Provide some insight – don't simply restate your results.