Name:_____

1 Probability

1. A thrilled birdwatcher suspects she has spotted an ivory-billed woodpecker. Skeptics contend that it was probably a common pileated woodpecker.

   - Ivory-billed woodpeckers are a class 6 species, defined as "definitely or probably extinct." The chance of seeing an ivory-billed woodpecker can be estimated as one in a million. The chance of seeing a pileated woodpecker can be estimated as one in 50.
   - The birdwatcher estimates her bird's length at 19 inches. That's on the large side for a pileated woodpecker, which has a 10% chance of being that size. An ivory-billed woodpecker has a 70% chance of being that size.

   What is the probability that the bird is an ivory-billed woodpecker?

   (Note: this answer is a very small number; you can leave it in terms of the component probabilities without multiplying, dividing, and adding it up.)

2. You are at the mouth of a cave full of treasure. You would like to enter the cave and get the treasure, but first you need to figure out if the dragon is inside.

   - If the dragon is home, there is a 70% chance you'll see some villagers on your way to the cave; if the dragon is out flying, there's a 90% chance they'll all stay home.
   - The probability of smoke coming from the cave is 80% if the dragon is home and 30% if the dragon is not home (the dragon's captive may be cooking something).
   - If the dragon went out the day before, the chance it will fly today is 50%. If it did not go out the day before, the chance it will today is 80%.
   - The prior probability of the dragon being out flying is 0.6.

   (a) Draw a dynamic Bayesian network corresponding to this problem, including the necessary conditional probability tables.

   (b) You have observed the cave for three days.

   $$\mathbf{e}_1 = smoky, \neg villagers$$
   $$\mathbf{e}_2 = smoky, villagers$$
   $$\mathbf{e}_3 = \neg smoky, villagers$$

   What is the probability that the dragon is home on Day 3?

2 Decision trees

Should you eat the food you just dropped on the floor?

Training data:

	Sticky	Kind of floor	Last piece	Eat it?
1	Yes	Kitchen	Yes	No
2	No	Kitchen	No	Yes
3	Yes	Bathroom	No	No
4	Yes	Bedroom	Yes	Yes
5	No	Bedroom	No	Yes
6	Yes	Bedroom	No	No
7	No	Bathroom	Yes	No
8	No	Kitchen	Yes	Yes

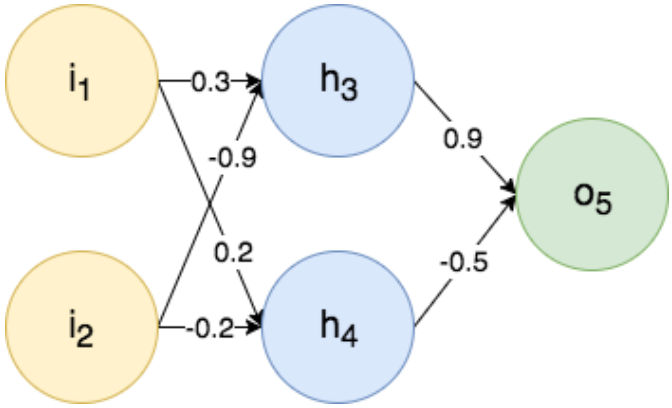
(a) What is the information gain from an initial split on each attribute (Sticky, Floor, or LastPiece)? Which one should a decision tree learning algorithm choose?

(b) Draw a decision tree representing the training data.

(c) How would the tree classify the following new instance? Trace the path of the classification through the decision tree.

	Sticky	Kind of floor	Last piece	Eat it?
	No	Kitchen	Yes	?

3 Neural networks



Suppose the training data for a neural network with the architecture depicted above is

$\mathbf{i_1}$	$\mathbf{i_2}$	$\mathbf{y}$
1	0	1
1	1	0

the activation function for the hidden and output nodes is $\max(x, 0)$

and the weights are initialized randomly with the values above

1. What is the output of the first forward pass for each training example?

2. What is the total error?
3. What is the gradient of the cost function $\frac{1}{2}(y - \hat{y})^2$ with respect to a_5?
4. What is the gradient of in_5 with respect to $w_{3,5}$?
5. What should the new value of each weight be after the first backpropagation pass?
6. What is the total error with the new weights?

4 General machine learning

Define (one short sentence each):

a) Overfitting

b) The curse of dimensionality

c) Regularization

d) Cross-validation

e) Gradient descent