Assignment: Probability, Supervised Learning, and Alignment

October 24, 2012

$Y \backslash X$	1	2	3	4
1	$\frac{1}{4}$	0	$\frac{1}{8}$	$\frac{1}{8}$
2	$\frac{1}{8}$	0	$\frac{1}{16}$	$\frac{1}{16}$
3	$\frac{1}{16}$	0	$\frac{1}{32}$	$\frac{1}{32}$
4	$\frac{1}{16}$	0	$\frac{1}{32}$	$\frac{1}{32}$

Figure 1: A joint distribution over the random variables X and Y.

1. **Probability, entropy, and mutual information.** Consider the joint distribution shown in Figure 1. Calculate the following from the joint distribution. (Please do this one completely on your own.) Note that many of the answers require more than a single number. Put answers on a separate sheet of paper.

   (a) $P((X+Y)=5)$

   (b) This is a weird one: $P(P(X,Y)=\frac{1}{16})$.

   (c) $P(X)$

   (d) $P(X|Y=3)$

   (e) $P(X=3|Y)$

   (f) $P(Y|X)$

   (g) $H(X)$. $H(\cdot)$ is the discrete entropy function. Please do this without a calculator.

   (h) $H(X,Y)$, the joint entropy of the distribution.

   (i) $I(X,Y)$, the mutual information between X and Y. Do this without a caculator.

   (j) Are X and Y independent?

2. **Color illusions**. Look at the color illusion in Figure 2. Explain what the illusion is, and whether this is a bug or a feature of the human visual system. Explain your answer in detail, including the reason that it would be helpful or harmful for the human visual system to see the world this way.

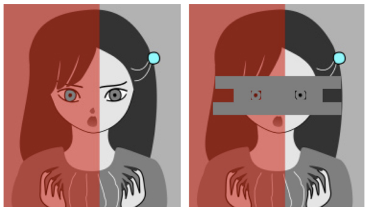

Figure 2: A color illusion.

3. **Bayes' rule**. Make up your own scenario in which a problem is solved with Bayes' rule. You should have two versions of the problem. In the first version, assume you are given the likelihoods for each class (as two separate tables), and the prior probabilities of each class as an additional table. In the second version of the problem, assume you are only given a sample of data from each class, and show how you would estimate the likelihoods and priors from the data. If you like, you can just estimate the likelihoods from the data, and make up the priors based upon your own reasoning.

   This problem will be graded on clarity and proper use of terminology in addition to whether the math is correct.

4. **Probability of error.** Nafi and Omar are playing a game with fair six-sided dice. Here's how it works. Nafi flips a coin that Omar cannot see. If it is heads, then Nafi rolls two dice X and Y and computes their sum and reports the number. If the coin is tails, then Nafi rolls a third die Z and reports $Z + 4$. Omar's job is to guess whether Nafi got heads or tails on the coin flip based on the number that Nafi reports. If Omar follows an optimal strategy, what is his probability of making an error on each trial of the experiment? Show the work you do to get to the result.

5. **Counting, Part A**. Consider 100 different horizontal shifts of an image, 100 vertical shifts, 50 different scalings, and 30 shears. How many of the following types of transformations can I make from these ingredients:

   (a) Affine transformations?

   (b) Rigid transformations?

   (c) Similarity transformations?

   (d) Translations?

6. **Counting, Part B.** Suppose Toshiko wants to use the following measurements as features into a face recognition classifier:

   - Distance between the eyes in millimeters, expressed as an integer.
   - Distance between bottom of nose and upper lip in mm, expressed as an integer.
   - Color of hair expressed as one of 10 different discrete values (blonde, brown, dark brown, black, red, etc.).
   - Width of nose in mm, expressed as an integer.

   Develop a reasonable upper bound on the *maximum* number of people that could be distinguished using these features.

7. **Color and light.** As we saw in one of the first lectures, the strength of the response of each type of cone cell c_i (where "i" can mean the red type, the blue type, or the green type of cone cells) can be written

$$c_i = \int s_i(\lambda) t(\lambda) \, d\lambda,$$

where s_i is the sensitivity of the ith cone type to each wavelength λ and t is the power of light at each wavelength λ.

Fred has two different flashlights, A and B, that he is pointing in your eye. (Fred can be annoying at times.) He claims that if the responses to the first flashlight are given by cone responses c_r^A, c_g^A, c_b^A, and the responses to the second flashlight are c_r^B, c_g^B, c_b^B, that he can predict the response of your cone cells when he shines both lights in your eye at the same time!

Show mathematically that the equation above implies that the cone cell responses are a linear function of the input lights. To do this, show that the sum of two lights $t_1(\lambda)$ and $t_2\lambda$ produce responses from the cone cells that are equal to the sum of the responses from the individual lights. That is, show that the response function is *linear* in the inputs. Also, give the response of the cone cells to the two flashlights simultaneously.