

CS294-082: Experimental Design for Machine Learning on Multimedia Data  
Fall 2020

Homework 1

To be discussed: September 11th, 2020

1) Machine learning from your experience

a) Draw a “box and arrow” typical workflow of a machine learning experiment. Include training and testing.

b) At which points of your workflow can you add quality controls?

2) Bit Arithmetic.

a) How many bits do you need to encode the integer 126?

b) How many bits do you need to encode 32.56? (technically and theoretically)

- Assume you have two positive integers of size  $n$  bits each: How many bits does the result of c) addition, d) subtraction, e) multiplication, of the two numbers maximally generate?

- Assume you have an  $8 \times 8$  matrix of 8 bit numbers. You now f) select the maximum number, g) calculate the average number of that matrix. How many bits do you need to store the result in each case?

- How many binary matrices of size  $8 \times 8$  are there in total?

3) Information Content

Assume you have a black and white (binary) image with a resolution of  $64 \times 64$  pixels.

What is the a) minimal and b) maximal information content in bits of that image?

c) How do the images from a) and b) look like?

4) Suppose 5 pairs of socks are in a drawer. How many socks do you have to minimally pick to guarantee that at least one pair is chosen?

5) Each of 15 red balls and 15 green balls is marked with an integer between 1 and 100 inclusive; no integer appears on more than one ball. The value of a pair of balls is the sum of the numbers on the balls. Show there are at least two pairs, consisting of one red and one green ball, with the same value. Show that this is not necessarily true if there are 13 balls of each color.

6) Estimate the memory required to implement a “20 Questions” game in a self-contained device (see <http://20q.net/>). The original game was implemented using a neural network trained on user input but, for this exercise, you can assume a decision tree.