

CS 541-A Artificial Intelligence: Mid-Term Exam

Instructor: Jie Shen

10/20/2020, 18:30 – 21:00 EST

Instructions:

- Open book exam, feel free to use any resource;
- Discussion is not permitted;
- Always give your answer and explain it (guaranteed 5 point for nonempty answer);
- 20 points per problem, totally 110 points ($20 * 5 + 10$).

0. Write down your name. (10 pts)

1. Choose a topic from the course that you are most interested in, and talk about your understanding.

2. Let X be a random variable. Suppose that $E[X] = 1$. What are the practical implications? If we further know that $\text{Var}[X] = 10$, what are the practical implications?

3. Given an image, suppose that an expert will present you the correct label with probability at least 0.99, and he charges 100 dollars. On the other side, it is possible to distribute the image to a pool of non-experts, each of which charges you 1 dollar but the probability that he returns the correct label is as low as p ($p > 0.5$). Therefore, when going with the second option we have to hire many workers and take majority vote. Give a sufficient condition on p such that the following two are fulfilled simultaneously:

- the quality of the label from majority voting is as good as the one from the expert;
- it costs less to hire these non-experts.

Note: The desired condition is like $f(p) \geq 0$ for some function f . You need to show what f is, but you do not need to calculate the value of p .

4. Random projection (RP) is a widely used tool for dimension reduction.

- What are the major advantages of RP?

- Give an example where RP fails to boost computational efficiency;
 - What do you think will happen if we use different types of random matrices, for example, discrete matrix for random projection? What are possible benefits and what are potential issues?
5. State the main idea of collaborative filtering. What are the possible drawbacks of its current formulation, and how to improve them?