# CS4851/6851 IDL Homework 5 April 6, 2022

Note: All coding problems to be submited with Github Link. Do not Upload the files/folder. Use git commands only.

Note: this is the distribution of questions:

1. Question 1 to Question 3: Required for everyone.
2. Question 4 to Question 5: Bonus question for both Graduate Students and Undergraduate Students

# Problem 1 (10 points)

For each of the following norms, explain what properties will they favor when used in reconstruction error: *L*0, *L*1, and *L*2

L0 : encourages weights to be zero and penalizes them when they’re different. Favors sparsity at a much greater scale.

L1: favors sparsity which is the property of having weights that are highly significant.

L2: favors small weights. As the difference between targets and predictions is squared, this norm discourages large weights that result in huge differences between predictions and targets.

# Problem 2 (30 points)

Given a set of contrast images with sharp geometric edges (e.g. photo-lithography masks for microprocessor manufacturing) write down a formulation for reconstruction error that would work best. Justify your choice.

I would use the reconstruction error that corresponds to L0 because it looks like we’re in a situation that requires a lot attention to small details. So, I would guess that weights would need to highly significant and also very close to reality.

# Problem 3 (20 points)

Given a set of images of wild life taken in their natural habitat write down a formulation for reconstruction error that would work best. Justify your choice.

sum from i space equals space 1 to m of open parentheses x to the power of i minus stack x to the power of i with overbrace on top close parentheses squared which corresponds to L2 norm. I think this choice is appropriate because we want to allow for small differences while punishing extremely large ones as there can be a variety of phenomes within the same family of animals.

Bonus for both undergraduates and gradu-

ates beyond this line.

# Problem 4 (40 points)

Given distributions p and q. If q is parameterized by *θ*, how would you choose the value for *θ* to make q closest to p among all possible q’s.

1. Write down formulation of how would you measure the closeness of *q* to *p*.
2. Explain what you would do to maximize this closeness (i.e. make *q* and *p* maximally close, or minimally different or divergent)

1

# Problem 5 (40 points)

Write a report on one of the following topics related to GANS:

1. InfoGAN https://arxiv.org/abs/1606.03657
2. CycleGAN

Summary of Info GAN: InfoGAN paper proposes a variation to GAN that makes the generative model a lot more efficient in learning salient features within the dataset. In the original GAN, the generator takes a noise input vector z, and attempts to map it into the true dataspace. InfoGAN adds another vector c that encodes the salient feature into the dataset. It argues that the relevance of an unsupervised model is its ability to be used for downstream tasks. The vector c records the salient features that may become relevant for other classification, regression, etc. tasks. In one experiment about classifying handwritten digits, the paper argues that one instance of c which was designated to capture drastic change in shape was used to classify handwritten digits, and it only achieved 5% error rate proving how relevant this implementation can be for downstream tasks.

2