CS294-082: Experimental Design for Machine Learning on Multimedia Data
Fall 2019

Homework 6

To be discussed: October 18th, 2019

1) Generalization vs Adversarial Examples

a) Show how definition 0 and definition 1 (from lecture 6 slides) are logically related (hint: use syntactic rules)

b) With a) in mind, how would you define adversarial examples using the English language?

2) Features and Generalization

Experiment with using features in the TFMeter app.

a) Describe what happens to the capacity requirement.

b) Would you still use features? Why or why not?

3) Debugging machine learning.

Use Shannon's communication model applied to the machine learning (for example, see MacKay's book Figure 40.1 or lecture 3, slide 18) to show that predicting the predictions of a machine learner (this is, accurately getting the results of a machine learner without asking the machine learner for the result) would be showing that P=NP.

4) With the solution of 3) in mind, outline general strategies to predicting adversarial examples.

5) "All models are wrong but some are useful"

Definition (classical physics): *An **observable** is a physical quantity that can be measured*. Assume a model *M* is perfect. Argue that *M* would not be an observable without assumptions or added external knowledge. For this exercise, it is easiest to pick a simple, (physical) example. No need for deep math.