

CS 7150: Deep Learning — Summer-Full 2020 — Paul Hand

Week 3 — Preparation Questions For Class

Due: Monday May 18, 2020 at 12:00 PM Eastern time via [Gradescope](#)

Name: [Put Your Name Here]

Collaborators: [Put Your Collaborators Here]

You may consult any and all resources in answering the questions. Your goal is to have answers that are ready to be shared with the class (or on a hypothetical job interview) as written.

Directions: Read the articles ‘[Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift](#)’ (original BN paper) and ‘[How Does Batch Normalization Help Optimization?](#)’ (Santurkar et al.) . Watch this [video](#) of Ian Goodfellow explaining batch normalization (3:46 - 13:16)

Question 1. *In the context of CNNs, what is Batch Normalization? In a BN layer, a variance is computed over a batch of images, but at test time there may be only a single image passed into the network. The variance of a quantity over only a single datapoint is undefined. How is this issue dealt with?*

Response:

Question 2. *What empirical benefits does batch normalization provide? Explain the experimental evidence for these benefits.*

Response:

Question 3. *In the original BN paper, what evidence is provided that batch normalization works because it reduces internal covariate shift? Is this evidence convincing?*

Response:

Question 4. *Explain the evidence that the Santurkar et al. paper uses to argue that BN’s performance is not explained by reducing internal covariate shift.*

Response:

Question 5. *What is Santurkar et al.’s explanation of BN’s performance. What experimental evidence is provided that batch normalization works because of this explanation? Is this evidence convincing?*

Response:

Question 6. *What theoretical results are made in the Santurkar et al. paper? Are these results convincing in explaining the behavior of batch normalization?*

Response:

Question 7. *What explanation does Ian Goodfellow provide for batch normalization. What reasoning is provided for believing it?*

Response:

Question 8. *What is the point of having an explanation for why batch normalization works?*

Response: