

## CS 7150: Deep Learning — Fall 2024 — Paul Hand

Day 3 — Preparation Questions For Class

Due: Tuesday 9/17/2024 at 2:50pm via Gradescope

Names: [Put The Names Of Your Group 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. Your answers should be as concise as possible. When asked to explain a figure, your response should have the following structure: provide context (state what experiment was being run / state what problem is being solved), state what has been plotted, remark on what we observe from the plots, and interpret the results.

Submit one document for your group and tag all group members. We recommend you use Overleaf for joint editing of this TeX document.

**Directions:** Read ‘Deep Learning’ (Three Giants Paper)

- Read the whole paper

**Question 1.** *What is representation learning? In what sense are deep learning methods considered to be representation learning?*

**Response:**

**Question 2.** *Explain Figure 1a. In your explanation, explain how it connects to the claim that ‘The hidden layers can be seen as distorting the input in a non-linear way so that categories become linearly separable by the last layer’. See the preamble above for comments on how to explain a figure.*

**Response:**

Context:

What is plotted:

What we observe:

Interpretation:

**Question 3.** *What is the selectivity-invariance dilemma and why is it challenging to overcome?*

**Response:**

**Question 4.** *Explain Figure 4.*

**Response:**

**Question 5.** *What do the authors think is the future of deep learning?*

**Response:**

**Directions:**

Read 'Understanding deep learning requires rethinking generalization.'

- Read only: Abstract, Sections 1.0, 1.1, 2.0, 2.1, 6

**Question 6.** *What is regularization? What is the explanation of why it is 'needed' to obtain good generalization performance?*

**Response:**

**Question 7.** *Explain Figure 1a.*

**Response:**

Context:

What is plotted:

What we observe:

Interpretation:

**Question 8.** *Explain Figures 1bc.*

**Response:**

Context:

What is plotted:

What we observe:

Interpretation:

**Question 9.** *In Section 6, the authors write: " The experiments we conducted emphasize that the effective capacity of several successful neural network architectures is large enough to shatter the training data. Consequently, these models are in principle rich enough to memorize the training data." What is meant by 'capacity of ... neural network architectures'? What is meant by 'shatter the training data'? What is meant by 'memorize the training data'?*

**Response:**

Capacity:

Shatter:

Memorize: