

Introduction to Deep Learning (Spring 2025)

Midterm

Date: April 8th, Tuesday 12:10pm - 3:10pm
Room: RWH105

- This exam is closed book i.e. no laptops, notes, textbooks, etc. during the exam. However, you may use one A4 sheet (front and back) of notes as reference. Calculator is also allowable.
- **For Q3-Q5, please organize your answer well and write them clearly.** We will give partial credit for **good and clear explanations** even though the answer is not fully correct.

The following are the sample questions for you to know in advance the question formats and the examining topics.
Note they are not the real exam questions!

1 True/False (20 points)

- Include 10 True/False questions and each is worth 2 points.
- These 10 questions involves basic knowledge about deep learning what you have learned so far. Review the lecture slide 1-9 to prepare for them.

Fill in the circle next to True or False, or fill in neither. No explanations are required. Correct answer is worth 2 points. To discourage guessing, incorrect answers are worth -1 points. Leaving a question blank will give 0 points.

Some sample questions:

1. A ReLU activation function can output negative values.
☐ True
☐ False
2. In order to normalize our data (i.e. subtract mean and divide by standard deviation), we typically compute the mean and standard deviation across the entire dataset before splitting the data into train/val/test splits.
☐ True
☐ False
3. Suppose we have trained a softmax classifier (with weights W) that achieves 100% accuracy on our dataset. If we change the weights to $2W$, the classifier will maintain the same accuracy while have a smaller loss (cross entropy loss without regularization).
☐ True
☐ False
4. Batch size refers to the number of training examples processed simultaneously before the parameters are updated.
☐ True
☐ False
5. Only deep neural networks with Sigmoid activation have vanishing gradients and exploding gradients problem.
☐ True
☐ False

2 Multiple Choices (20 points)

- Include 5 Multiple Choices questions and each worths 4 points.
- These 5 questions still involve basic knowledge about deep learning that you have learned so far. Review the lecture slides 1-9 to prepare for them.

Fill in the circle next to the letter(s) of your choice. No explanations are required. Choose ALL options that apply.

Each question is worth 4 points and the answer may contain one or more options. Selecting all of the correct options and none of the incorrect options will get full credits. For questions with multiple correct options, each incorrect or missing selection gets a 2-point deduction (up to 4 points).

1. We will be using gradient descent to optimize our losses in this question. Which of the following losses ($f(x)$) converge for the given initialization (x_0) and step size (step)? Select all that apply. Note: Our update rule is $x_{new} = x_{old} - step \times f'(x_{old})$.
 - ☐ A: $f(x) = |x|$ for $x_0 = 1$, step=2
 - ☐ B: $f(x) = |x|$ for $x_0 = 1$, step=2.01
 - ☐ C: $f(x) = x^2$ for $x_0 = 1$, step=1
 - ☐ D: $f(x) = x^2$ for $x_0 = 1$, step=0.99
2. Which of the following best describes the role of ReLU (Rectified Linear Unit) in a neural network?
 - ☐ A. It normalizes the outputs of each layer to zero mean and unit variance.
 - ☐ B. It sets negative activations to zero and preserves positive activations.
 - ☐ C. It applies a sigmoid function element-wise to the activations.
 - ☐ D. It is used only for output layers in classification tasks.
3. When training a CNN on image data, which statement is most accurate regarding the role of convolutional layers?
 - ☐ A. Each convolutional layer uses a completely different architecture from the others, with no parameter sharing.
 - ☐ B. Convolutional layers learn filters that capture local spatial patterns, and these filters are shared across the entire image.
 - ☐ C. Convolutional layers do not need activation functions because the convolutions themselves are nonlinear.
 - ☐ D. Convolutional layers can only handle black-and-white images, not color images.

3 Backpropagation (15 points)

This question will involve computational graph, backpropagation, gradient descent, ect. For sample questions, please refer to Homework2.

4 CNN and network training techniques (20 points)

This question will examine CNN, like its convolution and pooling mechanics, network structures, multi-input channel, etc., and some common training techniques, like dropout, batch normalization, so please review relevant slides. A sample question is as follows and a final exam question may includes multiple such questions.

A CNN consists of N layers. Each layer consists of a convolution with a 3×3 kernel, a sigmoid activation, and a max pooling operation with stride 2 in each dimension. There is no padding. The input image is 512×512 .

1) Compute the dimensions of the second and third layers. 2) What is the maximum value for N ?

5 Advanced Topics (25 points)

No sample question is given here because this question will involve some high-level concepts. These concepts are not taught directly in this class, but they are still based on the basic ideas and knowledge on deep learning, and other necessary information will be given directly in this question. You need to combine what you have learned and the given information to solve the problem. So do not need to worry about unfamiliar concepts shown in this question.