Deep Learning
Quiz 3

December 3, 2020

Instructor: Jia Xu

Student name: _____

Student ID: _____

Student email address: _____

- **Read these instructions carefully**
- Do *not* turn this page until you are instructed to do so
- Fill-in your personal info as indicated above
- You have 24 hours
- There are 10 questions.
- Each question worths the same (10 points)
- This is an open-book test.

good luck!

1 Question

Consider a Generative Adversarial Network (GAN) which successfully produces images of apples. Which of the following propositions is false?

A. The generator aims to learn the distribution of apple images.

B. The discriminator can be used to classify images as apple vs. non-apple.

C. After training the GAN, the discriminator loss eventually reaches a constant value.

D. The generator can produce unseen images of apples.

2 Question

After training a neural network, you observe a large gap between the training accuracy (100%) and the test accuracy (42%). Which of the following methods is commonly used to reduce this gap?

A. Generative Adversarial Networks

B. Dropout

C. Sigmoid activation

D. RMSprop optimizer

3 Question

Which of the following propositions are true about a CONV layer? (Check all that apply.)

A. The number of weights depends on the depth of the input volume.

B. The number of biases is equal to the number of filters.

C. The total number of parameters depends on the stride.

D. The total number of parameters depends on the padding.

4 Question

Which of the following statements is the best description of overfitting?

A. The network becomes specialized and learns the training set too well.

B. The network can predict the correct outputs for test examples which lie outside the range of the training examples.

C. The network does not contain enough adjustable parameters (e.g., hidden units) to find a good approximation to the unknown function which generated the training data.

5 Question

Auto-Encoders: auto-encoders are machines of the form

$\tilde{Y} = Decoder(W_d, Z)$

$Z = Encoder(W_e, Y)$

where W_e are the parameters of the encoder, and W_d the parameters of the decoder.

(a) (5 points) Given a training set $\{Y_i, i \in [1, P]\}$, write a possible loss function with which to train a sparse auto-encoder.

(b) (5 points) Assuming the decoder is linear Decoder $(W_d, Z) = W_d Z$, what constraint should we add to prevent the system from converging to a trivial and useless solution?

6 Question

Consider the following linear auto-encoder with 1 input and 1 output: $\tilde{x} = w_2 w_1 x$, trained with the squared reconstruction error:

$L(W) = \frac{1}{P} \sum_{i=1}^{P} \frac{1}{2}(x^i - w_2 w_1 x^i)^2$

The scalar training samples have variance 1.

(a) (5 points) What is the set of solutions (with 0 loss)?

(b) (5 points) Does the loss have a saddle point? Where?

7 Question

A neural network has been encrypted on a device, you can access neither its architecture, nor the values of its parameters. Is it possible to create an adversarial example to attack this network? Explain why.

8 Question

You want to solve a classification task. You first train your network on 20 samples. Training converges, but the training loss is very high. You then decide to train this network on 10,000 examples. Is your approach to fixing the problem correct? If yes, explain the most likely results of training with 10,000 examples. If not, give a solution to this problem.

9 Question

How does splitting a dataset into train, dev and test sets help identify overfitting?

10 Question

Give two benefits of using convolutional layers instead of fully connected ones for visual tasks.