EECS 182 Deep Neural Networks
Fall 2022 Anant Sahai

Homework 5

This homework is due on Saturday, October 15, 2022, at 10:59PM.

1. Understanding Dropout (Coding Question)

In this problem, you will fill out the dropout.ipynb notebook, which analyzes the effect of dropout in a simplified setting.

(A) Analytical gradient descent solution without dropout: **Fill out notebook section (A).**

(B) Empirical gradient descent solution without dropout. **Fill out notebook section (B).**

(C) Analytical gradient descent solution with dropout: **Fill out notebook section (C).**

(D) Empirical gradient descent solution without dropout. **Fill out notebook section (D)**

(E) Analyzing dropout's effect on the loss curve. **Run the cells in part (E), and explain the results. (Why is the loss curve so spiky, and why does the larger batch size help with matching the analytical solution?)**

(F) Refer back to the cells you ran in part (E). **Analyze how and why adding dropout changes the following: (a) How large were the final weights w_1 and w_2 compared to each other. (b) How large the contribution of each term (i.e. $10w_1$ + w_2) is to the final output. Why does this change occur?** (This does not need to be a formal math proof).

(G) (Optional) Sweeping over the dropout rate **Fill out notebook section (G).** You should see that as the dropout rate changes, w_1 and w_2 change smoothly, except for a discontinuity when dropout rates are 0. Explain this discontinuity.

(H) (Optional) Optimizing with Adam: Run the cells in part (H). **Does the solution change when you switch from SGD to Adam? Why or why not?**

(I) Dropout on real data: **Run the notebook cells in part (G), and report on how they affect the final performance.**

2. Inductive Bias and Systematic Experimentation

In this problem, you will fill out the EdgeDetection.ipynb notebook. We will learn 1) inductive bias of CNN and 2) systematic experimentation

(A) Overfitting Models to Small Dataset: **Fill out notebook section (Q1).**

  (a) Can you find any interesting patterns in the learned filters?

(B) Sweeping the Number of Training Images: **Fill out notebook section (Q2).**

  (a) Compare the learned kernels, untrained kernels, and edge-detector kernels. What do you observe?

  (b) We freeze the convolutional layer and train only final layer (classifier). In a high data regime, the performance of CNN initialized with edge detectors is worse than CNN initialized with random weights. Why do you think this happens?

(C) Checking the Training Procedures: **Fill out notebook section (Q3).**

(a) List every epochs that you trained the model. Final accuracy of CNN should be at least 90% for 20 images per class.
(b) Check the learned kernels. What do you observe?
(c) (optional) You might find that with the high number of epochs, validation loss of MLP is increasing whild validation accuracy increasing. How can we interpret this?
(d) (optional) Do hyperparameter tuning. And list the best hyperparameter setting that you found and report the final accuracy of CNN and MLP.
(e) How much more data is needed for MLP to get a competitive performance with CNN? Does MLP really generalize or memorize?

(D) Domain Shift between Training and Validation Set: **Fill out notebook section (Q4).**

(a) Why do you think the confusion matrix looks like this? Why CNN misclassifies the images with edge to the images with no edge? Why MLP misclassifies the images with vertical edge to the images with horizontal edge and vice versa? (Hint: Visualize some of the images in the training and validation set.)
(b) Why do you think MLP fails to learn the task while CNN can learn the task? (Hint: Think about the model architecture.)

(E) When CNN is Worse than MLP: **Fill out notebook section (Q5).**

(a) What do you observe? What is the reason that CNN is worse than MLP? (Hint: Think about the model architecture.)
(b) Assuming we are increasing kernel size of CNN. Does the validation accuracy increase or decrease? Why?
(c) How do the learned kernels look like? Explain why.

(F) Increasing the Number of Classes: **Fill out notebook section (Q6).**

3. Homework Process and Study Group

Citing sources and collaborators are an important part of life, including being a student!
We also want to understand what resources you find helpful and how much time homework is taking, so we can change things in the future if possible.

(a) **What sources (if any) did you use as you worked through the homework?**
(b) **If you worked with someone on this homework, who did you work with?**
List names and student ID's. (In case of homework party, you can also just describe the group.)
(c) **Roughly how many total hours did you work on this homework? Write it down here where you'll need to remember it for the self-grade form.**

Contributors:

- Olivia Watkins.
- Anant Sahai.
- Jake Austin.
- Suhong Moon.
- Kumar Krishna Agrawal.

