

Problem Set 2

The required weekly readings and lecture slides should be helpful in completing the assignment. You can find these on our [course website](#). For your submission, you must submit a PDF that has been typed. For mathematical derivations, if easier, you can insert pictures of your handwritten derivation into the final PDF.

1. Model Parameters versus Hyperparameters [3 points]:

- (a) In your own words, describe what is a model parameter and hyperparameter and how they differ.

2. Dataset Splits [4 points]:

- (a) In your own words, describe the motivation for creating a validation dataset and its purpose for developing deep learning models.
- (b) In your own words, describe what it means to create a stratified split when dividing your dataset into training, validation, and test splits and what is the motivation.

3. Overfitting versus Underfitting [2 points]:

- (a) If your model performs poorly on the training data and poorly on new examples, is the model overfitting or underfitting?
- (b) If your model performs well on the training data and generalizes poorly to new examples, is the model overfitting or underfitting?

4. Model Size [6 points]: Assume you are building a 5 layer neural network that takes as input a 32x32 matrix of integers, has 10 nodes per hidden layer, and has 100 nodes at the output layer.

- (a) If this network was fully connected, how many model parameters would it have? For full credit, you must indicate the number of model parameters between each pair of layers (e.g., input to hidden layer 1), the total number of weights in the network, and the total number of bias terms in the network.

- (b) If this network had convolutional layers with 5 3×3 filters per layer for all hidden layers, how many model parameters would it have? For full credit, you must indicate the number of model parameters between each pair of layers (e.g., input to hidden layer 1), the total number of weights in the network, and the total number of bias terms in the network.

5. Convolutional Neural Networks [10 points]:

- (a) In your own words, describe at least two advantages of using convolutional layers instead of fully connected layers in a neural network.
- (b) In your own words, describe at least two advantages of using pooling layers instead of fully connected layers in a neural network.
- (c) Compute and report the output that you would get from applying the filter on the input data, shown below, when using ‘SAME padding’. In order to get partial credit for the case that the output is incorrect, include your approach in your write-up.
- (d) Compute and report the output that you would get from applying the filter on the input data, shown below, when using 2×2 stride. In order to get partial credit for the case that the output is incorrect, include your approach in your write-up.

1	1	0	2
4	0	8	10
6	4	2	0
8	7	4	2
10	7	5	1

Data type 1: Input

0	0.5	0
0.5	1	0.5
0	0.5	0

Data type 2: Filter