

## Homework 4 Questions

### Document Instructions

- 5 questions [ $8 + 10 + 8 + 4 + 10 = 40$  points].
- Fill all your answers within the answer boxes, and **please do NOT remove the answer box outlines**.
- Include code, images, and equations where appropriate.
- To identify all places where your responses are expected, search for 'TODO'.
- Please make this document anonymous.

### Gradescope Instructions

- When you are finished, compile this document to a PDF and submit it directly to Gradescope.
- You will be required to assign the appropriate answers to the right pages on Gradescope. **Failure to assign pages correctly will lead to a deduction of 2 points per misaligned page (capped at a maximum 6 point deduction).**

**Q1: [8 points]**

- (a) **NEW:** In machine learning, there are two main types of metric errors. These errors appear when the magnitude of the model's error is averaged over different testing/training data sets. Describe the main difference between these two types of errors.

**OLD:** Describe these two types of machine learning.

- (i) **[1 point]** Bias

TODO: Your answer to (a) (i) here.

- (ii) **[1 point]** Variance

TODO: Your answer to (a) (ii) here.

- (iii) **[2 points]** Why is it especially dangerous to confuse one for the other?

TODO: Your answer to (a) (iii) here.

- (b) Define these terms in the context of evaluating a classifier:

- (i) **[1 point]** Overfitting

TODO: Your answer to (b) (i) here.

- (ii) **[1 point]** Underfitting

TODO: Your answer to (b) (ii) here.

- (c) **[2 points]** How do overfitting and underfitting relate to bias and variance?

TODO: Your answer to (c) here.

**Q2: [10 points]** Suppose we create a visual word dictionary using SIFT and k-means clustering for a scene recognition algorithm.

- (a) **[4 points]** What characteristics would be desirable in the dataset that you choose to use to build your visual dictionary? Consider [data bias trends](#) in popular machine learning datasets, the groups that these biases can harm, and how your choice in dataset could help. **[5-7 sentences]**

TODO: Your answer to (a) here.

- (b) **[4 points]** After choosing an appropriate dataset, examination of the SIFT features generated from our training database tells us that many are almost equidistant from two or more visual words.

- (i) **[2 points]** Why might this affect classification accuracy?

TODO: Your answer to (i) here.

- (ii) **[2 points]** Why shouldn't we pick out a new dataset that better matches our model?

TODO: Your answer to (ii) here.

- (c) **[2 points]** Given the situation, describe *two* methods to improve classification accuracy, and explain why they would help.  
*These can be for k-means, or otherwise.*

TODO: Your answer to (c) here.

**Q3: [8 points]** NEW: The bag of words representation handles the spatial layout of information in a way that can be an advantage or a disadvantage in different cases.

OLD: The way that the bag of words representation handles the spatial layout of visual information can be both an advantage and a disadvantage.

- (a) **[5 point]** NEW: How might we determine whether bag of words is a good model? What are good metrics for evaluating the bag of words model's effectiveness? Discuss both technical and nontechnical benchmarks for a "good" model. **[5-6 sentences]**

OLD: How might we determine whether bag of words is a good model?

TODO: Your answer to (a) here.

- (b) **[2 points]** NEW: Describe two example scenarios when dealing with visual information, one that illustrates an advantage of the bag of words representation, and another that shows a disadvantage of the bag of words representation.

OLD: Describe an example scenario for each of these cases.

TODO: Your answer to (b) here.

- (c) **[1 point]** Describe a modification or additional algorithm which might overcome the disadvantage.

TODO: Your answer to (c) here.

**Q4: [4 points]** Given a linear classifier such as SVM which separates two classes (binary decision), how might we use multiple linear classifiers to create a new classifier which separates  $k$  classes?

Below, we provide pseudocode for a linear classifier. It trains a model on a training set, and then classifies a new test example into one of two classes. Please edit the pseudo-code to convert this into a multi-class classifier.

*Hints:* See slides in supervised learning crash course deck, plus your own research. You can take either the one vs. all (or one vs. others) approach or the one vs. one approach in the slides; please declare which approach you take.

*More hints:* Be aware that:

1. The input labels in the multi-class case are different, and you will need to match the expected label input for the `train_linear_classifier` function
2. You need to make a new decision on how to aggregate or decide on the most confident prediction

*Note:* A more efficient software application would separate the classifier training and testing into two different functions so that the model could be reused without retraining. Feel free to ignore this for now.

TODO: Select the implementation you chose.

- |             |                          |
|-------------|--------------------------|
| One vs One  | <input type="checkbox"/> |
| One vs Many | <input type="checkbox"/> |

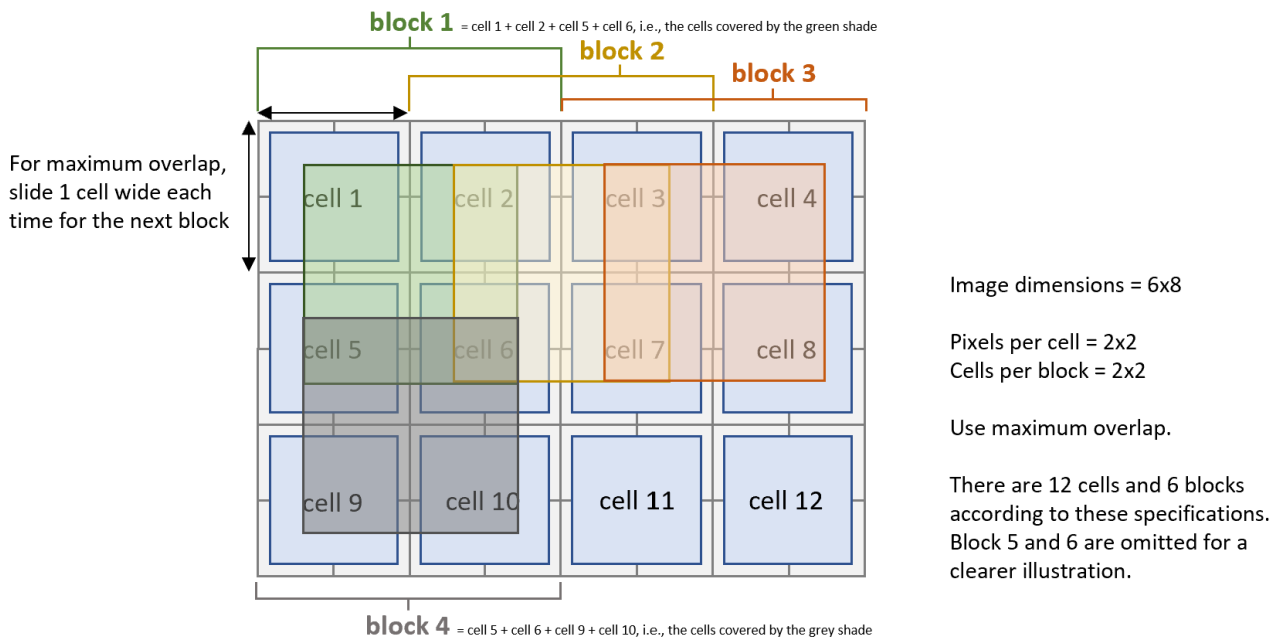
```
# Inputs
# train_feats: N x d matrix of N features each d descriptor long
# train_labels: N x 1 array containing values of either -1
#               (class 0) or 1 (class 1)
# test_feat: 1 x d image for which we wish to predict a label
#
# Outputs: -1 (class 0) or 1 (class 1)
#
# Inputs:
#   As before, except
#   train_labels: N x 1 array of class label integers from 0 to k-1
# Outputs:
#   A class label integer from 0 to k-1
#
# TODO: Turn this into a multi-class classifier for k classes.
def classify(train_feats, train_labels, test_feat)
    # Train classification hyperplane
    weights, bias = train_linear_classifier(train_feats, train_label)
    # Compute distance from hyperplane
    test_score = weights * test_feats + bias

    return 1 if test_score > 0 else -1
```

**Q5: [10 points]** In this homework, we will use a feature descriptor called HOG—‘Histogram of Oriented Gradients’. As its name implies, it works similarly to SIFT. In classification, we might extract HOG features across the entire image (not just at interest points).

HOG creates one feature descriptor per image ‘block’. Each block is split into ‘cells’ covering pixels. HOG outputs a 9-bin histogram of oriented gradients per cell. We append these together to obtain the feature descriptor for each block. As a result, if we have  $(z, z)$  cells per block, the feature descriptor for each block will be of size  $z \times z \times 9$ . NEW: In other words, computing HOG over the whole image produces a matrix where each row is a descriptor.

*Blocks can overlap as displayed in the diagram below.*



(a) **[6 points]** Given a  $72 \times 72$  image, calculate the number of cells, blocks, and final feature vector size that will occur when we extract HOG features over the whole image with the following parameters using maximum overlap between blocks.

(i) **[3 × 1 points]** Scenario 1: Pixels per cell =  $4 \times 4$ , cells per block =  $4 \times 4$

1. Number of cells:

TODO: Your answer to (a) (i) 1. here.

2. Number of blocks:

TODO: Your answer to (a) (i) 2. here.

3. Dimensions of feature vector for the whole image:

TODO: Your answer to (a) (i) 3. here.

(ii) [**3 × 1 points**] Scenario 2: Pixels per cell =  $8 \times 8$ , cells per block =  $2 \times 2$ .

1. Number of cells:

TODO: Your answer to (a) (ii) 1. here.

2. Number of blocks:

TODO: Your answer to (a) (ii) 2. here.

3. Dimensions of feature vector for the whole image:

TODO: Your answer to (a) (ii) 3. here.

(b) [**2 + 2 points**] When using HOG, the parameters such as pixels per cell and cells per block impact the resulting feature descriptor and so our performance on a classification task.

What are the pros and cons of the two parameter combinations? Which might you expect to have better performance?

**Feedback? (Optional)**

Please help us make the course better. If you have any feedback for this assignment, we'd love to hear it!