

knn

September 24, 2019

1 k-Nearest Neighbor (kNN) exercise

Complete and hand in this completed worksheet (including its outputs and any supporting code outside of the worksheet) with your assignment submission. For more details see the [assignments page](#) on the course website.

The kNN classifier consists of two stages:

- During training, the classifier takes the training data and simply remembers it
- During testing, kNN classifies every test image by comparing to all training images and transferring the labels of the k most similar training examples
- The value of k is cross-validated

In this exercise you will implement these steps and understand the basic Image Classification pipeline, cross-validation, and gain proficiency in writing efficient, vectorized code.

```
In [1]: # Run some setup code for this notebook.
        from __future__ import print_function

        import random
        import numpy as np
        from cs682.data_utils import load_CIFAR10
        import matplotlib.pyplot as plt

        # This is a bit of magic to make matplotlib figures appear inline in the notebook
        # rather than in a new window.
        %matplotlib inline
        plt.rcParams['figure.figsize'] = (10.0, 8.0) # set default size of plots
        plt.rcParams['image.interpolation'] = 'nearest'
        plt.rcParams['image.cmap'] = 'gray'

        # Some more magic so that the notebook will reload external python modules;
        # see http://stackoverflow.com/questions/1907993/autoreload-of-modules-in-ipython
        %load_ext autoreload
        %autoreload 2

In [2]: # Load the raw CIFAR-10 data.
        cifar10_dir = 'cs682/datasets/cifar-10-batches-py'

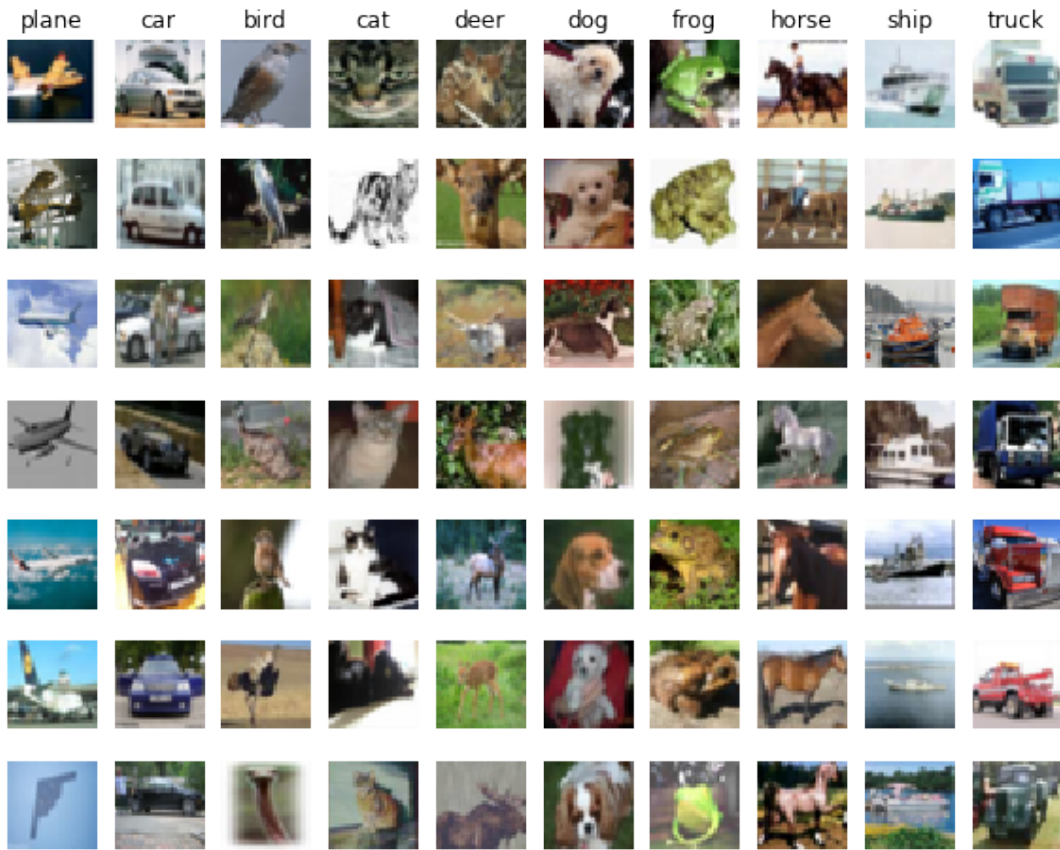
        # Cleaning up variables to prevent loading data multiple times (which may cause memory
        # issue)
        try:
            del X_train, y_train
            del X_test, y_test
            print('Clear previously loaded data.')
        except:
            pass

        X_train, y_train, X_test, y_test = load_CIFAR10(cifar10_dir)

        # As a sanity check, we print out the size of the training and test data.
        print('Training data shape: ', X_train.shape)
        print('Training labels shape: ', y_train.shape)
        print('Test data shape: ', X_test.shape)
        print('Test labels shape: ', y_test.shape)
```

```
Training data shape: (50000, 32, 32, 3)
Training labels shape: (50000,)
Test data shape: (10000, 32, 32, 3)
Test labels shape: (10000,)
```

```
In [3]: # Visualize some examples from the dataset.
# We show a few examples of training images from each class.
classes = ['plane', 'car', 'bird', 'cat', 'deer', 'dog', 'frog', 'horse', 'ship', 'truck']
num_classes = len(classes)
samples_per_class = 7
for y, cls in enumerate(classes):
    idxs = np.flatnonzero(y_train == y)
    idxs = np.random.choice(idxs, samples_per_class, replace=False)
    for i, idx in enumerate(idxs):
        plt_idx = i * num_classes + y + 1
        plt.subplot(samples_per_class, num_classes, plt_idx)
        plt.imshow(X_train[idx].astype('uint8'))
        plt.axis('off')
        if i == 0:
            plt.title(cls)
plt.show()
```



```
In [4]: # Subsample the data for more efficient code execution in this exercise
num_training = 5000
mask = list(range(num_training))
X_train = X_train[mask]
y_train = y_train[mask]

num_test = 500
mask = list(range(num_test))
X_test = X_test[mask]
y_test = y_test[mask]

In [5]: # Reshape the image data into rows
X_train = np.reshape(X_train, (X_train.shape[0], -1))
X_test = np.reshape(X_test, (X_test.shape[0], -1))
print(X_train.shape, y_train.shape, X_test.shape, y_test.shape)
```

(5000, 3072) (5000,) (500, 3072) (500,)

```
In [6]: from cs682.classifiers import KNearestNeighbor

# Create a kNN classifier instance.
# Remember that training a kNN classifier is a noop:
# the Classifier simply remembers the data and does no further processing
classifier = KNearestNeighbor()
classifier.train(X_train, y_train)
```

We would now like to classify the test data with the kNN classifier. Recall that we can break down this process into two steps:

1. First we must compute the distances between all test examples and all train examples.
2. Given these distances, for each test example we find the k nearest examples and have them vote for the label

Lets begin with computing the distance matrix between all training and test examples. For example, if there are **N_{tr}** training examples and **N_{te}** test examples, this stage should result in a **N_{te} x N_{tr}** matrix where each element (i,j) is the distance between the i-th test and j-th train example.

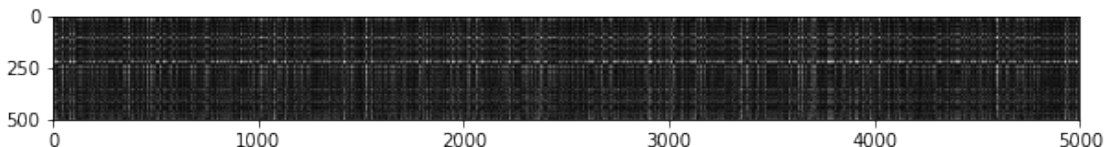
First, open `cs682/classifiers/k_nearest_neighbor.py` and implement the function `compute_distances_two_loops` that uses a (very inefficient) double loop over all pairs of (test, train) examples and computes the distance matrix one element at a time.

```
In [7]: # Open cs682/classifiers/k_nearest_neighbor.py and implement
# compute_distances_two_loops.
```

```
# Test your implementation:
dists = classifier.compute_distances_two_loops(X_test)
print(dists.shape)
```

(500, 5000)

```
In [8]: # We can visualize the distance matrix: each row is a single test example and
# its distances to training examples
plt.imshow(dists, interpolation='none')
plt.show()
```



Inline Question #1: Notice the structured patterns in the distance matrix, where some rows or columns are visible brighter. (Note that with the default color scheme black indicates low distances while white indicates high distances.)

- What in the data is the cause behind the distinctly bright rows?
- What causes the columns?

Your Answer: A bright row corresponds to a test image that is dissimilar from the majority of the images in the training set, where “dissimilar” is to be understood in the sense of a Euclidean distance, i.e. dissimilar images are “further apart”. This dissimilarity could be caused by e.g. an abnormally high concentration of a single color that doesn’t often occur in the training images. Dark rows, on the other hand, correspond to test images that are similar to the majority of training images.

The columns represent the same similarity and dissimilarity, only with the order of comparison reversed, i.e. a light column corresponds to a training image that is dissimilar from the majority of the test images, while a dark column corresponds to a training image that is similar to the majority of the test images.

```
In [9]: # Now implement the function predict_labels and run the code below:
# We use k = 1 (which is Nearest Neighbor).
y_test_pred = classifier.predict_labels(dists, k=1)

# Compute and print the fraction of correctly predicted examples
num_correct = np.sum(y_test_pred == y_test)
accuracy = float(num_correct) / num_test
print('Got %d / %d correct => accuracy: %f' % (num_correct, num_test, accuracy))
```

Got 137 / 500 correct => accuracy: 0.274000

You should expect to see approximately 27% accuracy. Now let's try out a larger k, say k = 5:

```
In [10]: y_test_pred = classifier.predict_labels(dists, k=5)
num_correct = np.sum(y_test_pred == y_test)
accuracy = float(num_correct) / num_test
print('Got %d / %d correct => accuracy: %f' % (num_correct, num_test, accuracy))
```

Got 139 / 500 correct => accuracy: 0.278000

You should expect to see a slightly better performance than with k = 1.

Inline Question 2 We can also use other distance metrics such as L1 distance. The performance of a Nearest Neighbor classifier that uses L1 distance will not change if (Select all that apply.): 1. The data is preprocessed by subtracting the mean. 2. The data is preprocessed by subtracting the mean and dividing by the standard deviation. 3. The coordinate axes for the data are rotated. 4. None of the above.

Your Answer: 1 and 2

Your explanation:

1. Subtracting the mean corresponds to shifting every image by the same constant amount. Because images are not being shifted relative to one another, the distance between them is invariant.
2. Subtracting the mean has no effect as described above. Dividing by the standard deviation corresponds to scaling the data by the same constant factor. This scaling may cause distances to increase or decrease depending on the magnitude of the scaling factor, but because everything is being scaled by the same amount, it will not change the relative ordering of distances between data, so the performance of a Nearest Neighbor classifier will not change.
3. Consider a unit square fixed at the origin and three data $X = \{a = (0,0), b = (1,0), c = (1,1)\}$ at three corners of the square. In the current frame, $|a - b| = 1$ and $|a - c| = 2$, therefore $|a - b| < |a - c|$. Now consider a $\pi/4$ rotation of the coordinate axes about the origin, so in the transformed frame the data becomes $X' = \{a' = (0,0), b' = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right), c' = (\sqrt{2}, 0)\}$, and the L1 distances between the data in the new frame are $|a' - b'| = \sqrt{2}$ and $|a' - c'| = \sqrt{2}$. Then in the transformed coordinate frame, $|a' - b'| = |a' - c'|$. Therefore the L1 norm is not invariant under rotations, and the performance of a Nearest Neighbor classifier will change because the relative ordering of data is not preserved.
4. See above

```
In [11]: # Now let's speed up distance matrix computation by using partial vectorization
# with one loop. Implement the function compute_distances_one_loop and run the
# code below:
dists_one = classifier.compute_distances_one_loop(X_test)

y_test_pred_one = classifier.predict_labels(dists_one, k=5)
num_correct_one = np.sum(y_test_pred_one == y_test)
accuracy_one = float(num_correct_one) / num_test
print('Got %d / %d correct => accuracy: %f' % (num_correct_one, num_test, accuracy_one))

# To ensure that our vectorized implementation is correct, we make sure that it
# agrees with the naive implementation. There are many ways to decide whether
# two matrices are similar; one of the simplest is the Frobenius norm. In case
# you haven't seen it before, the Frobenius norm of two matrices is the square
# root of the squared sum of differences of all elements; in other words, reshape
# the matrices into vectors and compute the Euclidean distance between them.
difference = np.linalg.norm(dists - dists_one, ord='fro')
print('Difference was: %f' % (difference, ))
if difference < 0.001:
    print('Good! The distance matrices are the same')
else:
    print('Uh-oh! The distance matrices are different')
```

Got 139 / 500 correct => accuracy: 0.278000
Difference was: 0.000000
Good! The distance matrices are the same

```
In [12]: # Now implement the fully vectorized version inside compute_distances_no_loops
# and run the code
dists_two = classifier.compute_distances_no_loops(X_test)

# check that the distance matrix agrees with the one we computed before:
difference = np.linalg.norm(dists - dists_two, ord='fro')
print('Difference was: %f' % (difference, ))
if difference < 0.001:
    print('Good! The distance matrices are the same')
else:
    print('Uh-oh! The distance matrices are different')
```

Difference was: 0.000000
Good! The distance matrices are the same

```
In [13]: # Let's compare how fast the implementations are
def time_function(f, *args):
    """
    Call a function f with args and return the time (in seconds) that it took to
    execute.
    """
    import time
    tic = time.time()
    f(*args)
    toc = time.time()
    return toc - tic

two_loop_time = time_function(classifier.compute_distances_two_loops, X_test)
print('Two loop version took %f seconds' % two_loop_time)

one_loop_time = time_function(classifier.compute_distances_one_loop, X_test)
print('One loop version took %f seconds' % one_loop_time)

no_loop_time = time_function(classifier.compute_distances_no_loops, X_test)
print('No loop version took %f seconds' % no_loop_time)

# you should see significantly faster performance with the fully vectorized
implementation
```

Two loop version took 45.679772 seconds
One loop version took 28.277893 seconds
No loop version took 0.384076 seconds

1.0.1 Cross-validation

We have implemented the k-Nearest Neighbor classifier but we set the value $k = 5$ arbitrarily. We will now determine the best value of this hyperparameter with cross-validation.

```
In [14]: num_folds = 5
k_choices = [1, 3, 5, 8, 10, 12, 15, 20, 50, 100]

#####
# TODO:
# Split up the training data into folds. After splitting, X_train_folds and
# y_train_folds should each be lists of length num_folds, where
# y_train_folds[i] is the label vector for the points in X_train_folds[i].
# Hint: Look up the numpy array_split function.
#####

X_train_folds = np.array_split(X_train, num_folds)
y_train_folds = np.array_split(y_train, num_folds)

#####
#                                     END OF YOUR CODE
#####

# A dictionary holding the accuracies for different values of k that we find
# when running cross-validation. After running cross-validation,
# k_to_accuracies[k] should be a list of length num_folds giving the different
# accuracy values that we found when using that value of k.
k_to_accuracies = {}
```

```

#####
# TODO:
# Perform k-fold cross validation to find the best value of k. For each
# possible value of k, run the k-nearest-neighbor algorithm num_folds times,
# where in each case you use all but one of the folds as training data and the
# last fold as a validation set. Store the accuracies for all fold and all
# values of k in the k_to_accuracies dictionary.
#####

for k in k_choices:
    # array to store training accuracy for each validation fold
    accuracies = [None]*num_folds

    for i in range(num_folds):
        # current validation fold
        x_validation_fold = X_train_folds[i]
        y_validation_fold = y_train_folds[i]

        # concatenate all other folds into one array for classifier training
        x_training_fold = np.concatenate([X_train_folds[j] for j in range(num_folds) if
j is not i])
        y_training_fold = np.concatenate([y_train_folds[j] for j in range(num_folds) if
j is not i])

        # train new classifier & predict labels of validation fold
        classifier.train(x_training_fold, y_training_fold)
        y_validation_pred = classifier.predict(x_validation_fold, k)

        # compute and accuracy
        num_correct = np.sum(y_validation_pred == y_validation_fold)
        accuracy = float(num_correct) / y_validation_fold.shape[0]
        accuracies[i] = accuracy

    k_to_accuracies[k] = accuracies

#####
#                                     END OF YOUR CODE                                     #
#####

# Print out the computed accuracies
for k in sorted(k_to_accuracies):
    for accuracy in k_to_accuracies[k]:
        print('k = %d, accuracy = %f' % (k, accuracy))
    print('-----')
    print('k = %d, mean accuracy = %f' % (k, np.mean(k_to_accuracies[k])))
    print('-----')

k = 1, accuracy = 0.263000
k = 1, accuracy = 0.257000
k = 1, accuracy = 0.264000
k = 1, accuracy = 0.278000
k = 1, accuracy = 0.266000
-----
k = 1, mean accuracy = 0.265600
-----
k = 3, accuracy = 0.239000
k = 3, accuracy = 0.249000
k = 3, accuracy = 0.240000
k = 3, accuracy = 0.266000
k = 3, accuracy = 0.254000
-----
k = 3, mean accuracy = 0.249600
-----
k = 5, accuracy = 0.248000
k = 5, accuracy = 0.266000
k = 5, accuracy = 0.280000
k = 5, accuracy = 0.292000
k = 5, accuracy = 0.280000
-----
k = 5, mean accuracy = 0.273200
-----
k = 8, accuracy = 0.262000
k = 8, accuracy = 0.282000
k = 8, accuracy = 0.273000
k = 8, accuracy = 0.290000

```

```

k = 8, accuracy = 0.273000
-----
k = 8, mean accuracy = 0.276000
-----
k = 10, accuracy = 0.265000
k = 10, accuracy = 0.296000
k = 10, accuracy = 0.276000
k = 10, accuracy = 0.284000
k = 10, accuracy = 0.280000
-----
k = 10, mean accuracy = 0.280200
-----
k = 12, accuracy = 0.260000
k = 12, accuracy = 0.295000
k = 12, accuracy = 0.279000
k = 12, accuracy = 0.283000
k = 12, accuracy = 0.280000
-----
k = 12, mean accuracy = 0.279400
-----
k = 15, accuracy = 0.252000
k = 15, accuracy = 0.289000
k = 15, accuracy = 0.278000
k = 15, accuracy = 0.282000
k = 15, accuracy = 0.274000
-----
k = 15, mean accuracy = 0.275000
-----
k = 20, accuracy = 0.270000
k = 20, accuracy = 0.279000
k = 20, accuracy = 0.279000
k = 20, accuracy = 0.282000
k = 20, accuracy = 0.285000
-----
k = 20, mean accuracy = 0.279000
-----
k = 50, accuracy = 0.271000
k = 50, accuracy = 0.288000
k = 50, accuracy = 0.278000
k = 50, accuracy = 0.269000
k = 50, accuracy = 0.266000
-----
k = 50, mean accuracy = 0.274400
-----
k = 100, accuracy = 0.256000
k = 100, accuracy = 0.270000
k = 100, accuracy = 0.263000
k = 100, accuracy = 0.256000
k = 100, accuracy = 0.263000
-----
k = 100, mean accuracy = 0.261600
-----

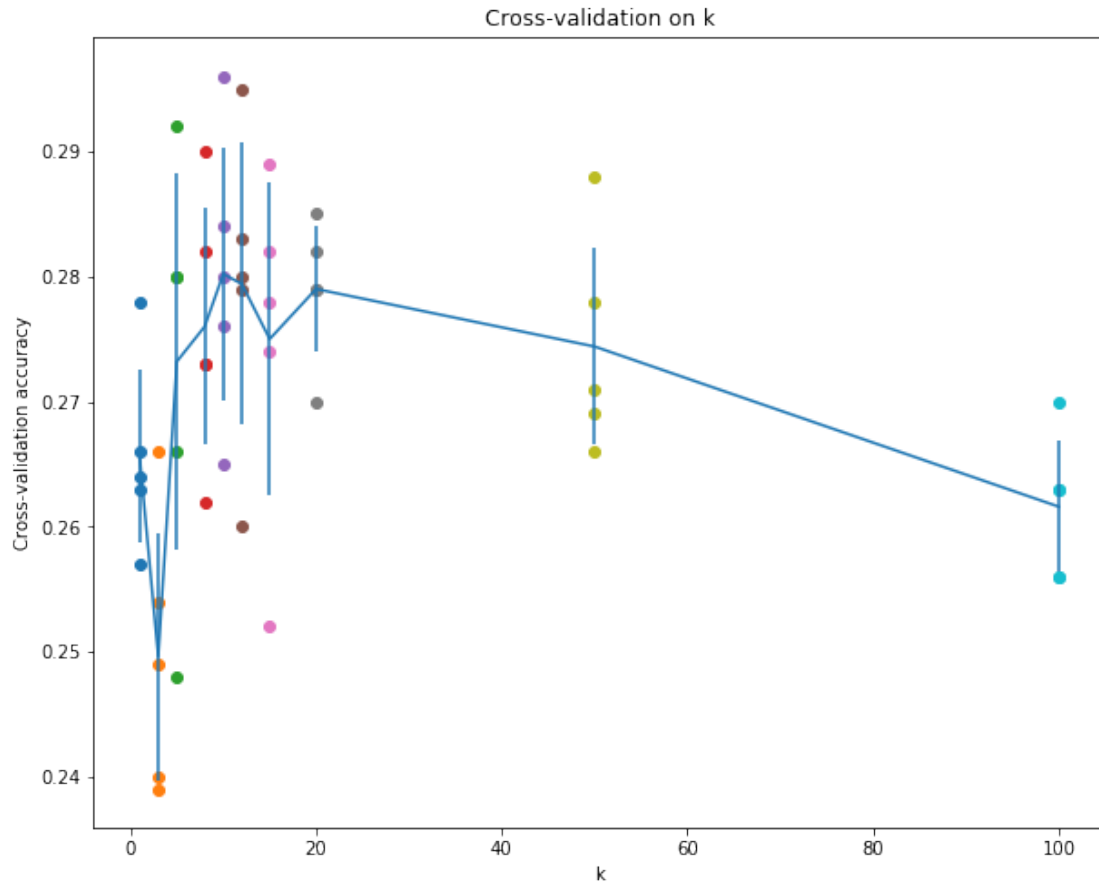
```

```

In [15]: # plot the raw observations
         for k in k_choices:
             accuracies = k_to_accuracies[k]
             plt.scatter([k] * len(accuracies), accuracies)

         # plot the trend line with error bars that correspond to standard deviation
         accuracies_mean = np.array([np.mean(v) for k,v in sorted(k_to_accuracies.items())])
         accuracies_std = np.array([np.std(v) for k,v in sorted(k_to_accuracies.items())])
         plt.errorbar(k_choices, accuracies_mean, yerr=accuracies_std)
         plt.title('Cross-validation on k')
         plt.xlabel('k')
         plt.ylabel('Cross-validation accuracy')
         plt.show()

```



```
In [16]: # Based on the cross-validation results above, choose the best value for k,
# retrain the classifier using all the training data, and test it on the test
# data. You should be able to get above 28% accuracy on the test data.
best_k = 10

classifier = KNearestNeighbor()
classifier.train(X_train, y_train)
y_test_pred = classifier.predict(X_test, k=best_k)

# Compute and display the accuracy
num_correct = np.sum(y_test_pred == y_test)
accuracy = float(num_correct) / num_test
print('Got %d / %d correct => accuracy: %f' % (num_correct, num_test, accuracy))

Got 141 / 500 correct => accuracy: 0.282000
```

Inline Question 3 Which of the following statements about k -Nearest Neighbor (k -NN) are true in a classification setting, and for all k ? Select all that apply. 1. The training error of a 1-NN will always be better than that of 5-NN. 2. The test error of a 1-NN will always be better than that of a 5-NN. 3. The decision boundary of the k -NN classifier is linear. 4. The time needed to classify a test example with the k -NN classifier grows with the size of the training set. 5. None of the above.

Your Answer: 4

Your explanation:

1. If the training data consists of binary class labels, and the data for each class is grouped in a tight cluster with the clusters far apart from each other, then (assuming each cluster contains at least five data) 1-NN and 5-NN will have the same training error. A more accurate statement would be, "The training error of a 1-NN will always be at least as good as that of 5-NN."

2. If the training data consists of 100 data with binary class labels $\{0, 1\}$, where 99 are labeled “class 1” and a single data is labeled “class 0,” and the data is uniformly distributed, then a random test example labeled “class 0” is likely to be misclassified by a 1-NN, but will be correctly classified by a 5-NN, resulting in a better test error for the 5-NN than 1-NN.
3. If the training data consists of 100 Euclidean points with binary class labels $\{0, 1\}$, where 99 points are labeled “class 1” and a single point is labeled “class 0,” then the decision boundary for 1-NN is a circle centered at the single “class 0” point with radius equal to half the distance between the point and its closest neighbor.
4. Classification requires comparing a test example to the stored training data, therefore classification time increases linearly with the size of the training set.
5. See above.

In []:

SVM

September 24, 2019

1 Multiclass Support Vector Machine exercise

Complete and hand in this completed worksheet (including its outputs and any supporting code outside of the worksheet) with your assignment submission. For more details see the [assignments page](#) on the course website.

In this exercise you will:

- implement a fully-vectorized **loss function** for the SVM
- implement the fully-vectorized expression for its **analytic gradient**
- **check your implementation** using numerical gradient
- use a validation set to **tune the learning rate and regularization strength**
- **optimize** the loss function with **SGD**
- **visualize** the final learned weights

In [2]: # Run some setup code for this notebook.

```
import random
import numpy as np
from cs682.data_utils import load_CIFAR10
import matplotlib.pyplot as plt

from __future__ import print_function

# This is a bit of magic to make matplotlib figures appear inline in the
# notebook rather than in a new window.
%matplotlib inline
plt.rcParams['figure.figsize'] = (10.0, 8.0) # set default size of plots
plt.rcParams['image.interpolation'] = 'nearest'
plt.rcParams['image.cmap'] = 'gray'

# Some more magic so that the notebook will reload external python modules;
# see http://stackoverflow.com/questions/1907993/autoreload-of-modules-in-ipython
%load_ext autoreload
%autoreload 2
```

1.1 CIFAR-10 Data Loading and Preprocessing

In [3]: # Load the raw CIFAR-10 data.

```
cifar10_dir = 'cs682/datasets/cifar-10-batches-py'

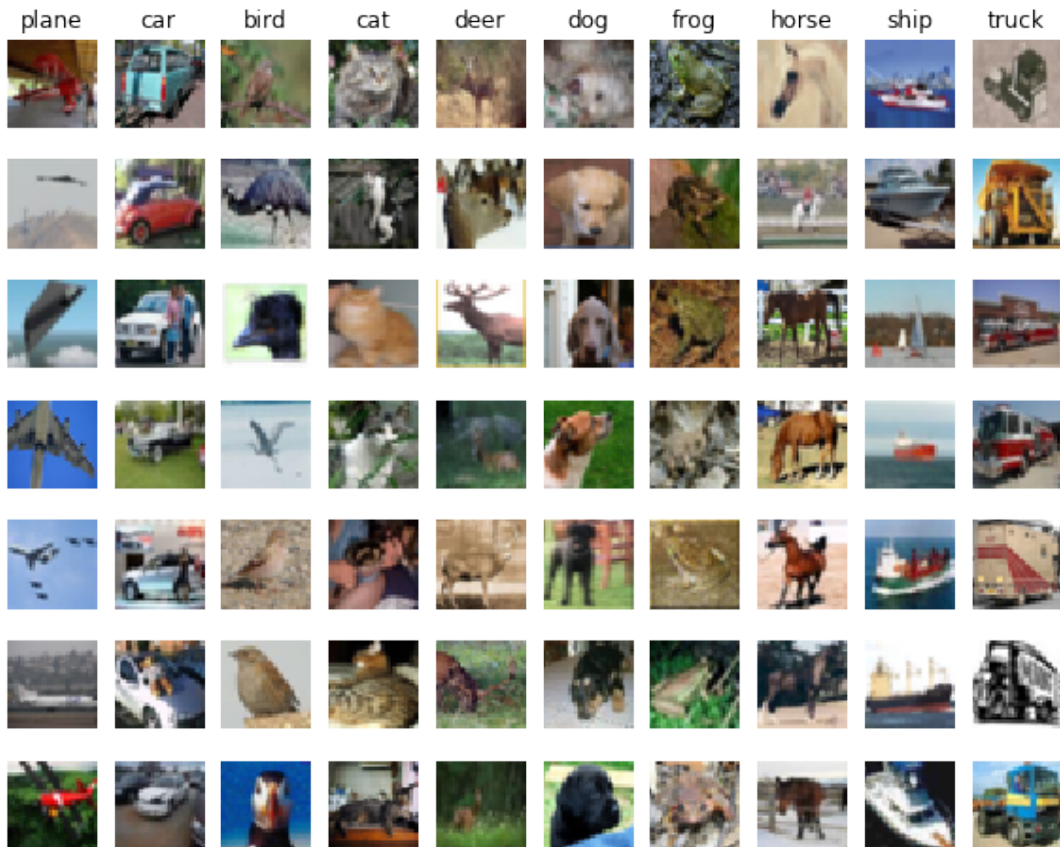
# Cleaning up variables to prevent loading data multiple times (which may cause memory
# issue)
try:
    del X_train, y_train
    del X_test, y_test
    print('Clear previously loaded data.')
except:
    pass

X_train, y_train, X_test, y_test = load_CIFAR10(cifar10_dir)

# As a sanity check, we print out the size of the training and test data.
print('Training data shape: ', X_train.shape)
print('Training labels shape: ', y_train.shape)
print('Test data shape: ', X_test.shape)
print('Test labels shape: ', y_test.shape)
```

Training data shape: (50000, 32, 32, 3)
 Training labels shape: (50000,)
 Test data shape: (10000, 32, 32, 3)
 Test labels shape: (10000,)

```
In [5]: # Visualize some examples from the dataset.
# We show a few examples of training images from each class.
classes = ['plane', 'car', 'bird', 'cat', 'deer', 'dog', 'frog', 'horse', 'ship', 'truck']
num_classes = len(classes)
samples_per_class = 7
for y, cls in enumerate(classes):
    idxs = np.flatnonzero(y_train == y)
    idxs = np.random.choice(idxs, samples_per_class, replace=False)
    for i, idx in enumerate(idxs):
        plt_idx = i * num_classes + y + 1
        plt.subplot(samples_per_class, num_classes, plt_idx)
        plt.imshow(X_train[idx].astype('uint8'))
        plt.axis('off')
        if i == 0:
            plt.title(cls)
plt.show()
```



```
In [6]: # Split the data into train, val, and test sets. In addition we will
# create a small development set as a subset of the training data;
# we can use this for development so our code runs faster.
num_training = 49000
num_validation = 1000
num_test = 1000
num_dev = 500

# Our validation set will be num_validation points from the original
```

```

# training set.
mask = range(num_training, num_training + num_validation)
X_val = X_train[mask]
y_val = y_train[mask]

# Our training set will be the first num_train points from the original
# training set.
mask = range(num_training)
X_train = X_train[mask]
y_train = y_train[mask]

# We will also make a development set, which is a small subset of
# the training set.
mask = np.random.choice(num_training, num_dev, replace=False)
X_dev = X_train[mask]
y_dev = y_train[mask]

# We use the first num_test points of the original test set as our
# test set.
mask = range(num_test)
X_test = X_test[mask]
y_test = y_test[mask]

print('Train data shape: ', X_train.shape)
print('Train labels shape: ', y_train.shape)
print('Validation data shape: ', X_val.shape)
print('Validation labels shape: ', y_val.shape)
print('Test data shape: ', X_test.shape)
print('Test labels shape: ', y_test.shape)

```

```

Train data shape: (49000, 32, 32, 3)
Train labels shape: (49000,)
Validation data shape: (1000, 32, 32, 3)
Validation labels shape: (1000,)
Test data shape: (1000, 32, 32, 3)
Test labels shape: (1000,)

```

```

In [7]: # Preprocessing: reshape the image data into rows
X_train = np.reshape(X_train, (X_train.shape[0], -1))
X_val = np.reshape(X_val, (X_val.shape[0], -1))
X_test = np.reshape(X_test, (X_test.shape[0], -1))
X_dev = np.reshape(X_dev, (X_dev.shape[0], -1))

# As a sanity check, print out the shapes of the data
print('Training data shape: ', X_train.shape)
print('Validation data shape: ', X_val.shape)
print('Test data shape: ', X_test.shape)
print('dev data shape: ', X_dev.shape)
print('y dev data shape:', y_dev.shape)

```

```

Training data shape: (49000, 3072)
Validation data shape: (1000, 3072)
Test data shape: (1000, 3072)
dev data shape: (500, 3072)
y dev data shape: (500,)

```

```

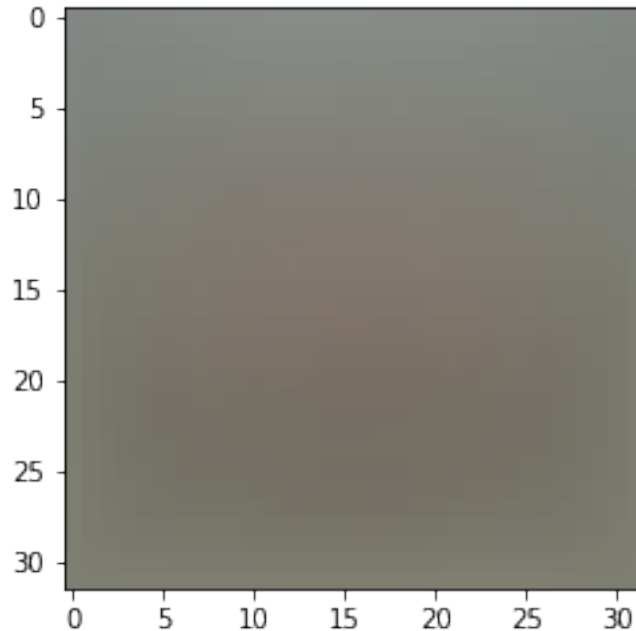
In [8]: # Preprocessing: subtract the mean image
# first: compute the image mean based on the training data
mean_image = np.mean(X_train, axis=0)
print(mean_image[:10]) # print a few of the elements
plt.figure(figsize=(4,4))
plt.imshow(mean_image.reshape((32,32,3)).astype('uint8')) # visualize the mean image
plt.show()

```

```

[130.64189796 135.98173469 132.47391837 130.05569388 135.34804082
 131.75402041 130.96055102 136.14328571 132.47636735 131.48467347]

```



```
In [9]: # second: subtract the mean image from train and test data
X_train -= mean_image
X_val -= mean_image
X_test -= mean_image
X_dev -= mean_image
```

1.2 NOTE: ONLY RUN THIS CELL ONCE; RUNNING MULTIPLE TIMES WILL RESULT IN SUBSEQUENT APPENDING OF BIAS DIMENSIONS, BREAKING EVERYTHING

```
In [10]: # third: append the bias dimension of ones (i.e. bias trick) so that our SVM
# only has to worry about optimizing a single weight matrix W.
X_train = np.hstack([X_train, np.ones((X_train.shape[0], 1))])
X_val = np.hstack([X_val, np.ones((X_val.shape[0], 1))])
X_test = np.hstack([X_test, np.ones((X_test.shape[0], 1))])
X_dev = np.hstack([X_dev, np.ones((X_dev.shape[0], 1))])
```

```
In [11]: # get the new shape without appending more bias dimensions
print(X_train.shape, X_val.shape, X_test.shape, X_dev.shape)
```

```
(49000, 3073) (1000, 3073) (1000, 3073) (500, 3073)
```

1.3 SVM Classifier

Your code for this section will all be written inside `cs682/classifiers/linear_svm.py`.

As you can see, we have prefilled the function `svm_loss_naive` which uses for loops to evaluate the multiclass SVM loss function.

```
In [12]: # Evaluate the naive implementation of the loss we provided for you:
from cs682.classifiers.linear_svm import svm_loss_naive
import time

# generate a random SVM weight matrix of small numbers
W = np.random.randn(3073, 10) * 0.0001

loss, grad = svm_loss_naive(W, X_dev, y_dev, 0.000005)
print('loss: %f' % (loss, ))
```

```
loss: 8.883228
```

The `grad` returned from the function above is right now all zero. Derive and implement the gradient for the SVM cost function and implement it inline inside the function `svm_loss_naive`. You will find it helpful to interleave your new code inside the existing function.

To check that you have correctly implemented the gradient correctly, you can numerically estimate the gradient of the loss function and compare the numeric estimate to the gradient that you computed. We have provided code that does this for you:

```
In [13]: # Once you've implemented the gradient, recompute it with the code below
# and gradient check it with the function we provided for you

# Compute the loss and its gradient at W.
loss, grad = svm_loss_naive(W, X_dev, y_dev, 0.0)

# Numerically compute the gradient along several randomly chosen dimensions, and
# compare them with your analytically computed gradient. The numbers should match
# almost exactly along all dimensions.
from cs682.gradient_check import grad_check_sparse
f = lambda w: svm_loss_naive(w, X_dev, y_dev, 0.0)[0]
grad_numerical = grad_check_sparse(f, W, grad)

# do the gradient check once again with regularization turned on
# you didn't forget the regularization gradient did you?
loss, grad = svm_loss_naive(W, X_dev, y_dev, 5e1)
f = lambda w: svm_loss_naive(w, X_dev, y_dev, 5e1)[0]
grad_numerical = grad_check_sparse(f, W, grad)

numerical: -11.574642 analytic: -11.574642, relative error: 1.832390e-11
numerical: -1.964842 analytic: -1.967475, relative error: 6.694744e-04
numerical: -6.904052 analytic: -6.904052, relative error: 3.623706e-11
numerical: -6.872089 analytic: -6.872089, relative error: 2.821836e-11
numerical: 7.126284 analytic: 7.126284, relative error: 1.593585e-11
numerical: 13.396511 analytic: 13.396511, relative error: 8.799963e-12
numerical: -8.702400 analytic: -8.757140, relative error: 3.135262e-03
numerical: -8.629595 analytic: -8.624555, relative error: 2.920535e-04
numerical: 0.634884 analytic: 0.634884, relative error: 3.117467e-10
numerical: -8.038393 analytic: -8.038393, relative error: 4.548692e-11
numerical: -60.236968 analytic: -60.236968, relative error: 5.977341e-12
numerical: 2.787166 analytic: 2.787166, relative error: 3.311749e-12
numerical: -5.733529 analytic: -5.733529, relative error: 1.367394e-11
numerical: -1.344836 analytic: -1.344836, relative error: 1.697072e-10
numerical: -5.413733 analytic: -5.413733, relative error: 1.580819e-11
numerical: -3.145625 analytic: -3.145625, relative error: 1.317116e-11
numerical: 4.225663 analytic: 4.225663, relative error: 7.162545e-11
numerical: 2.788418 analytic: 2.788418, relative error: 1.992738e-10
numerical: 13.837059 analytic: 13.765609, relative error: 2.588520e-03
numerical: -1.576365 analytic: -1.576365, relative error: 3.282397e-10
```

1.3.1 Inline Question 1:

It is possible that once in a while a dimension in the gradcheck will not match exactly. What could such a discrepancy be caused by? Is it a reason for concern? What is a simple example in one dimension where a gradient check could fail? How would change the margin affect of the frequency of this happening? Hint: the SVM loss function is not strictly speaking differentiable

Your Answer: One reason that the analytic and numerical gradients could not match exactly is because the step-size when computing the numerical gradient is too large to capture fluctuations in the function value. However, numerical gradients are accurate to order $\mathcal{O}(h^2)$, so these inaccuracies shouldn't be much of a concern.

This can also occur at points where the function is not differentiable (such as in the case of the svm loss function). This is also not a concern in the case of the svm loss because, although the derivative is not defined at the hinge point, the subderivative is, so any numerical approximations will still remain relatively close.

```
In [14]: # Next implement the function svm_loss_vectorized; for now only compute the loss;
# we will implement the gradient in a moment.
```

```

tic = time.time()
loss_naive, grad_naive = svm_loss_naive(W, X_dev, y_dev, 0.000005)
toc = time.time()
print('Naive loss: %e computed in %fs' % (loss_naive, toc - tic))

from cs682.classifiers.linear_svm import svm_loss_vectorized
tic = time.time()
loss_vectorized, _ = svm_loss_vectorized(W, X_dev, y_dev, 0.000005)
toc = time.time()
print('Vectorized loss: %e computed in %fs' % (loss_vectorized, toc - tic))

# The losses should match but your vectorized implementation should be much faster.
print('difference: %f' % (loss_naive - loss_vectorized))

```

Naive loss: 8.883228e+00 computed in 0.319426s
Vectorized loss: 8.883228e+00 computed in 0.007208s
difference: 0.000000

In [15]: *# Complete the implementation of svm_loss_vectorized, and compute the gradient
of the loss function in a vectorized way.*

```

# The naive implementation and the vectorized implementation should match, but  

# the vectorized version should still be much faster.
tic = time.time()
_, grad_naive = svm_loss_naive(W, X_dev, y_dev, 0.000005)
toc = time.time()
print('Naive loss and gradient: computed in %fs' % (toc - tic))

tic = time.time()
_, grad_vectorized = svm_loss_vectorized(W, X_dev, y_dev, 0.000005)
toc = time.time()
print('Vectorized loss and gradient: computed in %fs' % (toc - tic))

# The loss is a single number, so it is easy to compare the values computed  

# by the two implementations. The gradient on the other hand is a matrix, so  

# we use the Frobenius norm to compare them.
difference = np.linalg.norm(grad_naive - grad_vectorized, ord='fro')
print('difference: %f' % difference)

```

Naive loss and gradient: computed in 0.352965s
Vectorized loss and gradient: computed in 0.011503s
difference: 0.000000

1.3.2 Stochastic Gradient Descent

We now have vectorized and efficient expressions for the loss, the gradient and our gradient matches the numerical gradient. We are therefore ready to do SGD to minimize the loss.

In [16]: *# In the file linear_classifier.py, implement SGD in the function
LinearClassifier.train() and then run it with the code below.*

```

from cs682.classifiers import LinearSVM
svm = LinearSVM()
tic = time.time()
loss_hist = svm.train(X_train, y_train, learning_rate=1e-7, reg=2.5e4,
                      num_iters=2500, verbose=True)
toc = time.time()
print('That took %fs' % (toc - tic))

```

```

iteration 0 / 2500: loss 26.345077
iteration 100 / 2500: loss 7.467807
iteration 200 / 2500: loss 6.150301
iteration 300 / 2500: loss 5.228824
iteration 400 / 2500: loss 4.780265
iteration 500 / 2500: loss 4.682518
iteration 600 / 2500: loss 4.314449
iteration 700 / 2500: loss 4.511159
iteration 800 / 2500: loss 5.036806
iteration 900 / 2500: loss 4.601152
iteration 1000 / 2500: loss 5.009104
iteration 1100 / 2500: loss 4.904058
iteration 1200 / 2500: loss 4.891886
iteration 1300 / 2500: loss 4.524778
iteration 1400 / 2500: loss 4.591935
iteration 1500 / 2500: loss 4.734364

```

```

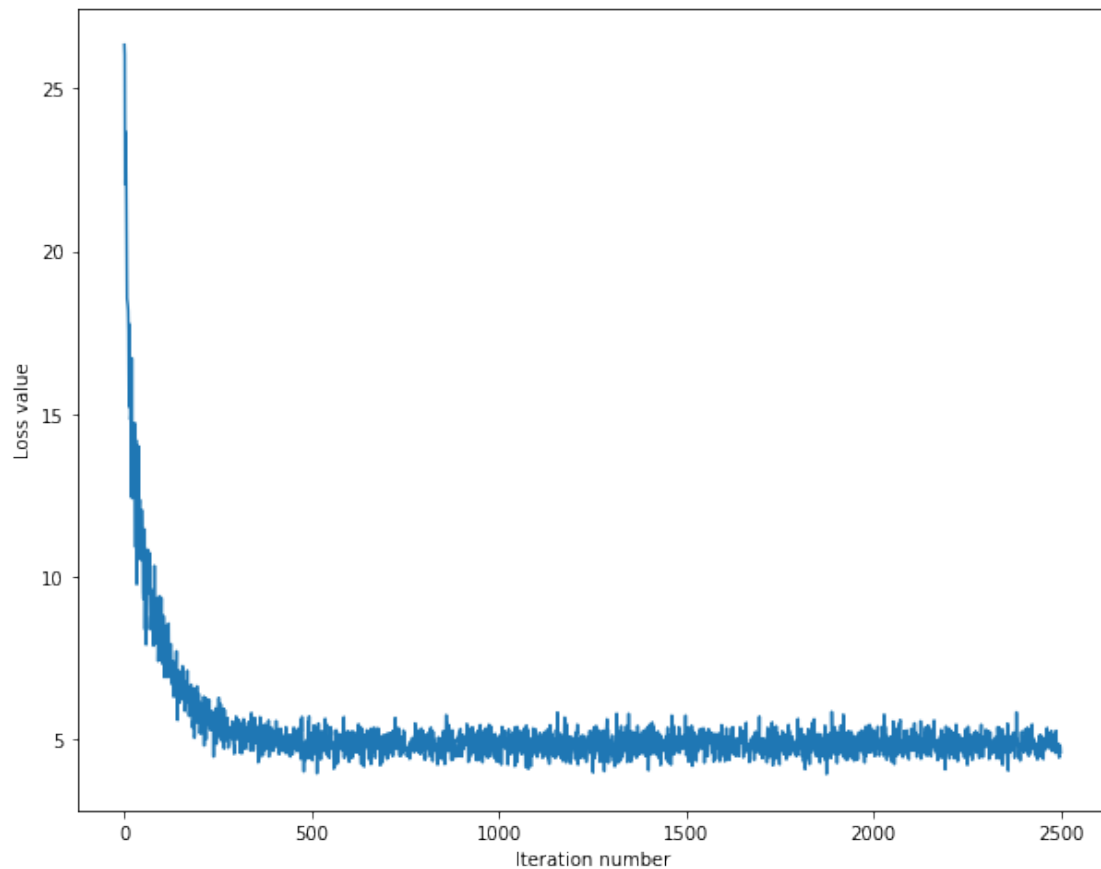
iteration 1600 / 2500: loss 4.925021
iteration 1700 / 2500: loss 4.831816
iteration 1800 / 2500: loss 4.691417
iteration 1900 / 2500: loss 5.244576
iteration 2000 / 2500: loss 5.007126
iteration 2100 / 2500: loss 5.326535
iteration 2200 / 2500: loss 5.185470
iteration 2300 / 2500: loss 4.812868
iteration 2400 / 2500: loss 4.864761
That took 13.767878s

```

```

In [17]: # A useful debugging strategy is to plot the loss as a function of
# iteration number:
plt.plot(loss_hist)
plt.xlabel('Iteration number')
plt.ylabel('Loss value')
plt.show()

```



```

In [18]: # Write the LinearSVM.predict function and evaluate the performance on both the
# training and validation set
y_train_pred = svm.predict(X_train)
print('training accuracy: %f' % (np.mean(y_train == y_train_pred), ))
y_val_pred = svm.predict(X_val)
print('validation accuracy: %f' % (np.mean(y_val == y_val_pred), ))

```

```

training accuracy: 0.370000
validation accuracy: 0.391000

```

```

In [25]: # Use the validation set to tune hyperparameters (regularization strength and
# learning rate). You should experiment with different ranges for the learning

```



```

# rates and regularization strengths; if you are careful you should be able to
# get a classification accuracy of about 0.4 on the validation set.
learning_rates = [5e-8, 1e-7, 2.5e-7, 5e-7, 1e-6, 2.5e-6, 5e-6]
regularization_strengths = [1e3, 2.5e3, 5e3, 1e4, 2.5e4, 5e4, 1e5]

# results is dictionary mapping tuples of the form
# (learning_rate, regularization_strength) to tuples of the form
# (training_accuracy, validation_accuracy). The accuracy is simply the fraction
# of data points that are correctly classified.
results = {}
best_val = -1 # The highest validation accuracy that we have seen so far.
best_svm = None # The LinearSVM object that achieved the highest validation rate.

```

```

#####
# TODO: #
# Write code that chooses the best hyperparameters by tuning on the validation #
# set. For each combination of hyperparameters, train a linear SVM on the #
# training set, compute its accuracy on the training and validation sets, and #
# store these numbers in the results dictionary. In addition, store the best #
# validation accuracy in best_val and the LinearSVM object that achieves this #
# accuracy in best_svm. #
# #
# Hint: You should use a small value for num_iters as you develop your #
# validation code so that the SVMs don't take much time to train; once you are #
# confident that your validation code works, you should rerun the validation #
# code with a larger value for num_iters. #
#####

```

```

for rate in learning_rates:
    for strength in regularization_strengths:
        svm_val = LinearSVM()
        loss_hist = svm_val.train(X_train,
                                   y_train,
                                   learning_rate=rate,
                                   reg=strength,
                                   num_iters=2500,
                                   verbose=False)

        yt_pred = svm_val.predict(X_train)
        yt_acc = np.mean(y_train == yt_pred)
        yv_pred = svm_val.predict(X_val)
        yv_acc = np.mean(y_val == yv_pred)

        print('rate = %e, strength = %e, train acc = %f, val acc = %f' % (rate,
                                   strength, yt_acc, yv_acc))

        results[(rate, strength)] = (yt_acc, yv_acc)

        if yv_acc > best_val:
            best_val = yv_acc
            best_svm = svm_val

```

```

#####
#                               END OF YOUR CODE                               #
#####

```

```

print('best validation accuracy achieved during cross-validation: %f' % best_val)

```

```

rate = 5.000000e-08, strength = 1.000000e+03, train acc = 0.309959, val acc = 0.318000
rate = 5.000000e-08, strength = 2.500000e+03, train acc = 0.328347, val acc = 0.356000
rate = 5.000000e-08, strength = 5.000000e+03, train acc = 0.365388, val acc = 0.363000
rate = 5.000000e-08, strength = 1.000000e+04, train acc = 0.381204, val acc = 0.383000
rate = 5.000000e-08, strength = 2.500000e+04, train acc = 0.368755, val acc = 0.379000
rate = 5.000000e-08, strength = 5.000000e+04, train acc = 0.362714, val acc = 0.382000
rate = 5.000000e-08, strength = 1.000000e+05, train acc = 0.346531, val acc = 0.348000
rate = 1.000000e-07, strength = 1.000000e+03, train acc = 0.351898, val acc = 0.344000
rate = 1.000000e-07, strength = 2.500000e+03, train acc = 0.384184, val acc = 0.383000
rate = 1.000000e-07, strength = 5.000000e+03, train acc = 0.391694, val acc = 0.387000
rate = 1.000000e-07, strength = 1.000000e+04, train acc = 0.388776, val acc = 0.385000
rate = 1.000000e-07, strength = 2.500000e+04, train acc = 0.367633, val acc = 0.383000
rate = 1.000000e-07, strength = 5.000000e+04, train acc = 0.360184, val acc = 0.363000
rate = 1.000000e-07, strength = 1.000000e+05, train acc = 0.330286, val acc = 0.345000
rate = 2.500000e-07, strength = 1.000000e+03, train acc = 0.397918, val acc = 0.379000
rate = 2.500000e-07, strength = 2.500000e+03, train acc = 0.388265, val acc = 0.385000
rate = 2.500000e-07, strength = 5.000000e+03, train acc = 0.372184, val acc = 0.378000
rate = 2.500000e-07, strength = 1.000000e+04, train acc = 0.379612, val acc = 0.379000
rate = 2.500000e-07, strength = 2.500000e+04, train acc = 0.358122, val acc = 0.375000

```

```

rate = 2.500000e-07, strength = 5.000000e+04, train acc = 0.345755, val acc = 0.363000
rate = 2.500000e-07, strength = 1.000000e+05, train acc = 0.339980, val acc = 0.363000
rate = 5.000000e-07, strength = 1.000000e+03, train acc = 0.391490, val acc = 0.401000
rate = 5.000000e-07, strength = 2.500000e+03, train acc = 0.378571, val acc = 0.369000
rate = 5.000000e-07, strength = 5.000000e+03, train acc = 0.362061, val acc = 0.349000
rate = 5.000000e-07, strength = 1.000000e+04, train acc = 0.363265, val acc = 0.359000
rate = 5.000000e-07, strength = 2.500000e+04, train acc = 0.329714, val acc = 0.354000
rate = 5.000000e-07, strength = 5.000000e+04, train acc = 0.335714, val acc = 0.334000
rate = 5.000000e-07, strength = 1.000000e+05, train acc = 0.284510, val acc = 0.283000
rate = 1.000000e-06, strength = 1.000000e+03, train acc = 0.357694, val acc = 0.345000
rate = 1.000000e-06, strength = 2.500000e+03, train acc = 0.347531, val acc = 0.362000
rate = 1.000000e-06, strength = 5.000000e+03, train acc = 0.333184, val acc = 0.345000
rate = 1.000000e-06, strength = 1.000000e+04, train acc = 0.321082, val acc = 0.322000
rate = 1.000000e-06, strength = 2.500000e+04, train acc = 0.298327, val acc = 0.277000
rate = 1.000000e-06, strength = 5.000000e+04, train acc = 0.295000, val acc = 0.292000
rate = 1.000000e-06, strength = 1.000000e+05, train acc = 0.245653, val acc = 0.253000
rate = 2.500000e-06, strength = 1.000000e+03, train acc = 0.316143, val acc = 0.315000
rate = 2.500000e-06, strength = 2.500000e+03, train acc = 0.287714, val acc = 0.280000
rate = 2.500000e-06, strength = 5.000000e+03, train acc = 0.285122, val acc = 0.290000
rate = 2.500000e-06, strength = 1.000000e+04, train acc = 0.252837, val acc = 0.278000
rate = 2.500000e-06, strength = 2.500000e+04, train acc = 0.234735, val acc = 0.246000
rate = 2.500000e-06, strength = 5.000000e+04, train acc = 0.222041, val acc = 0.232000
rate = 2.500000e-06, strength = 1.000000e+05, train acc = 0.187429, val acc = 0.202000
rate = 5.000000e-06, strength = 1.000000e+03, train acc = 0.278000, val acc = 0.276000
rate = 5.000000e-06, strength = 2.500000e+03, train acc = 0.266694, val acc = 0.264000
rate = 5.000000e-06, strength = 5.000000e+03, train acc = 0.275347, val acc = 0.294000
rate = 5.000000e-06, strength = 1.000000e+04, train acc = 0.232429, val acc = 0.237000
rate = 5.000000e-06, strength = 2.500000e+04, train acc = 0.201673, val acc = 0.199000
rate = 5.000000e-06, strength = 5.000000e+04, train acc = 0.155102, val acc = 0.140000
rate = 5.000000e-06, strength = 1.000000e+05, train acc = 0.142408, val acc = 0.143000
best validation accuracy achieved during cross-validation: 0.401000

```

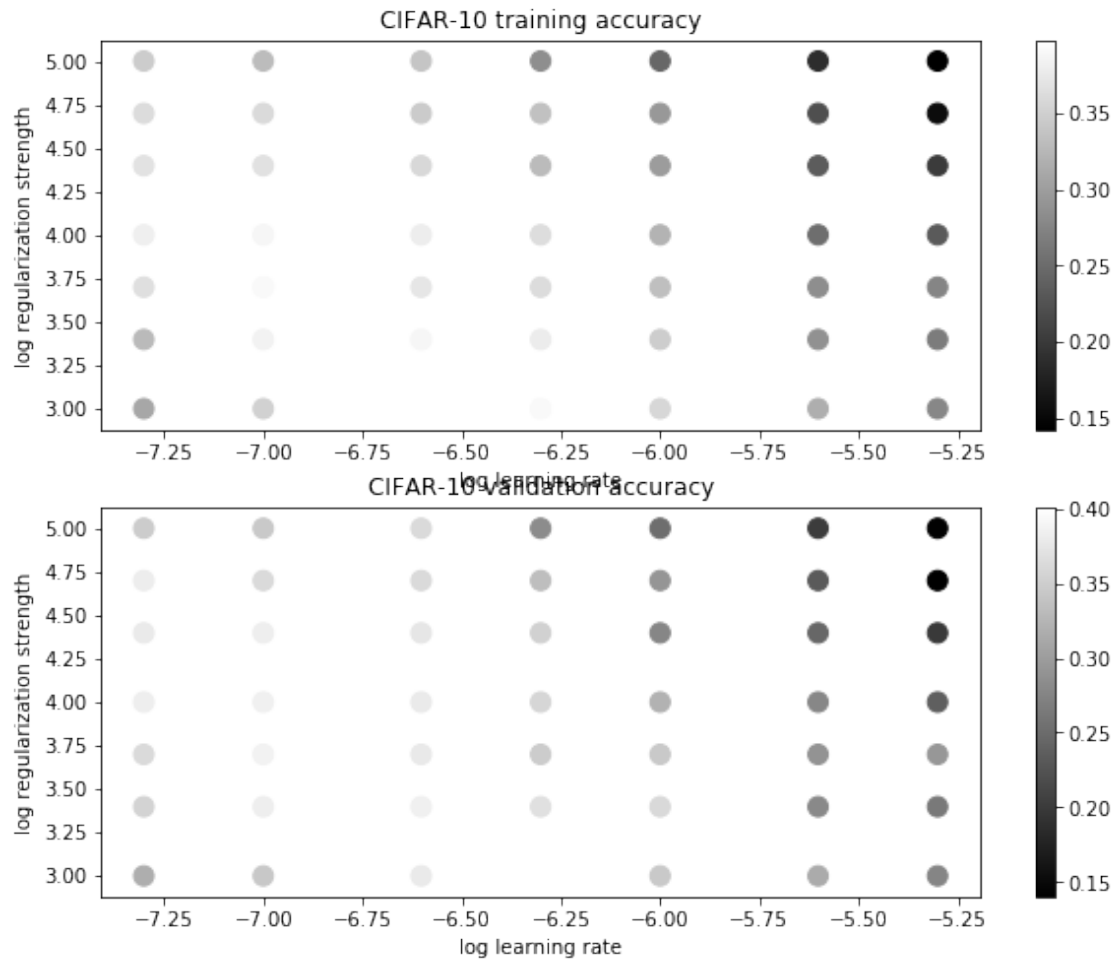
```

In [26]: # Visualize the cross-validation results
import math
x_scatter = [math.log10(x[0]) for x in results]
y_scatter = [math.log10(x[1]) for x in results]

# plot training accuracy
marker_size = 100
colors = [results[x][0] for x in results]
plt.subplot(2, 1, 1)
plt.scatter(x_scatter, y_scatter, marker_size, c=colors)
plt.colorbar()
plt.xlabel('log learning rate')
plt.ylabel('log regularization strength')
plt.title('CIFAR-10 training accuracy')

# plot validation accuracy
colors = [results[x][1] for x in results] # default size of markers is 20
plt.subplot(2, 1, 2)
plt.scatter(x_scatter, y_scatter, marker_size, c=colors)
plt.colorbar()
plt.xlabel('log learning rate')
plt.ylabel('log regularization strength')
plt.title('CIFAR-10 validation accuracy')
plt.show()

```

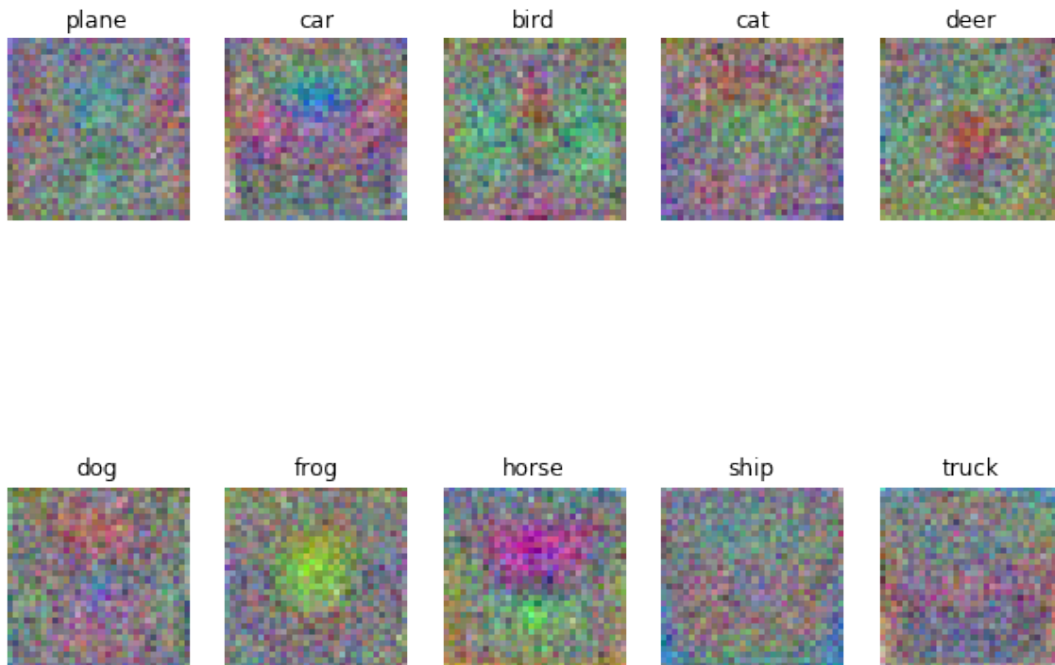


```
In [27]: # Evaluate the best svm on test set
y_test_pred = best_svm.predict(X_test)
test_accuracy = np.mean(y_test == y_test_pred)
print('linear SVM on raw pixels final test set accuracy: %f' % test_accuracy)
```

linear SVM on raw pixels final test set accuracy: 0.366000

```
In [28]: # Visualize the learned weights for each class.
# Depending on your choice of learning rate and regularization strength, these may
# or may not be nice to look at.
w = best_svm.W[:-1,:] # strip out the bias
w = w.reshape(32, 32, 3, 10)
w_min, w_max = np.min(w), np.max(w)
classes = ['plane', 'car', 'bird', 'cat', 'deer', 'dog', 'frog', 'horse', 'ship',
'truck']
for i in range(10):
    plt.subplot(2, 5, i + 1)

    # Rescale the weights to be between 0 and 255
    wimg = 255.0 * (w[:, :, :, i].squeeze() - w_min) / (w_max - w_min)
    plt.imshow(wimg.astype('uint8'))
    plt.axis('off')
    plt.title(classes[i])
```



1.3.3 Inline question 2:

Describe what your visualized SVM weights look like, and offer a brief explanation for why they look the way that they do.

Your answer: For the most part the weights look like noise, and although most of them have some discernable structure to them, it's hard to make out what that structure actually is. A few of the weights stand out: The car weights seem to have what could be a windshield in the center of the image, and doors or body panels on either side if you imagine looking at the car from an angle; The frog weights look like a green blob, which is not the worst description of what a frog looks like; If you squint, the horse weights have an area in the middle that's shaped like a two-headed horse facing in either direction. Each of these suggests the training data contains images of the intended target class (car, frog, horse) positioned in a number of different orientations, and that the SVM learned these weights by "averaging" over all of them, in some sense

softmax

September 24, 2019

1 Softmax exercise

Complete and hand in this completed worksheet (including its outputs and any supporting code outside of the worksheet) with your assignment submission. For more details see the [assignments page](#) on the course website.

This exercise is analogous to the SVM exercise. You will:

- implement a fully-vectorized **loss function** for the Softmax classifier
- implement the fully-vectorized expression for its **analytic gradient**
- **check your implementation** with numerical gradient
- use a validation set to **tune the learning rate and regularization strength**
- **optimize** the loss function with **SGD**
- **visualize** the final learned weights

```
In [1]: import random
import numpy as np
from cs682.data_utils import load_CIFAR10
import matplotlib.pyplot as plt

from __future__ import print_function

%matplotlib inline
plt.rcParams['figure.figsize'] = (10.0, 8.0) # set default size of plots
plt.rcParams['image.interpolation'] = 'nearest'
plt.rcParams['image.cmap'] = 'gray'

# for auto-reloading external modules
# see http://stackoverflow.com/questions/1907993/autoreload-of-modules-in-ipython
%load_ext autoreload
%autoreload 2

In [2]: def get_CIFAR10_data(num_training=49000, num_validation=1000, num_test=1000,
num_dev=500):
    """
    Load the CIFAR-10 dataset from disk and perform preprocessing to prepare
    it for the linear classifier. These are the same steps as we used for the
    SVM, but condensed to a single function.
    """
    # Load the raw CIFAR-10 data
    cifar10_dir = 'cs682/datasets/cifar-10-batches-py'

    X_train, y_train, X_test, y_test = load_CIFAR10(cifar10_dir)

    # subsample the data
    mask = list(range(num_training, num_training + num_validation))
    X_val = X_train[mask]
    y_val = y_train[mask]
    mask = list(range(num_training))
    X_train = X_train[mask]
    y_train = y_train[mask]
    mask = list(range(num_test))
    X_test = X_test[mask]
    y_test = y_test[mask]
    mask = np.random.choice(num_training, num_dev, replace=False)
    X_dev = X_train[mask]
    y_dev = y_train[mask]

    # Preprocessing: reshape the image data into rows
```

```

X_train = np.reshape(X_train, (X_train.shape[0], -1))
X_val = np.reshape(X_val, (X_val.shape[0], -1))
X_test = np.reshape(X_test, (X_test.shape[0], -1))
X_dev = np.reshape(X_dev, (X_dev.shape[0], -1))

# Normalize the data: subtract the mean image
mean_image = np.mean(X_train, axis = 0)
X_train -= mean_image
X_val -= mean_image
X_test -= mean_image
X_dev -= mean_image

# add bias dimension and transform into columns
X_train = np.hstack([X_train, np.ones((X_train.shape[0], 1))])
X_val = np.hstack([X_val, np.ones((X_val.shape[0], 1))])
X_test = np.hstack([X_test, np.ones((X_test.shape[0], 1))])
X_dev = np.hstack([X_dev, np.ones((X_dev.shape[0], 1))])

return X_train, y_train, X_val, y_val, X_test, y_test, X_dev, y_dev

# Cleaning up variables to prevent loading data multiple times (which may cause memory
issue)
try:
    del X_train, y_train
    del X_test, y_test
    print('Clear previously loaded data.')
except:
    pass

# Invoke the above function to get our data.
X_train, y_train, X_val, y_val, X_test, y_test, X_dev, y_dev = get_CIFAR10_data()
print('Train data shape: ', X_train.shape)
print('Train labels shape: ', y_train.shape)
print('Validation data shape: ', X_val.shape)
print('Validation labels shape: ', y_val.shape)
print('Test data shape: ', X_test.shape)
print('Test labels shape: ', y_test.shape)
print('dev data shape: ', X_dev.shape)
print('dev labels shape: ', y_dev.shape)

```

```

Train data shape: (49000, 3073)
Train labels shape: (49000,)
Validation data shape: (1000, 3073)
Validation labels shape: (1000,)
Test data shape: (1000, 3073)
Test labels shape: (1000,)
dev data shape: (500, 3073)
dev labels shape: (500,)

```

1.1 Softmax Classifier

Your code for this section will all be written inside `cs682/classifiers/softmax.py`.

```

In [3]: # First implement the naive softmax loss function with nested loops.
# Open the file cs682/classifiers/softmax.py and implement the
# softmax_loss_naive function.

from cs682.classifiers.softmax import softmax_loss_naive
import time

# Generate a random softmax weight matrix and use it to compute the loss.
W = np.random.randn(3073, 10) * 0.0001
loss, grad = softmax_loss_naive(W, X_dev, y_dev, 0.0)

# As a rough sanity check, our loss should be something close to -log(0.1).
print('loss: %f' % loss)
print('sanity check: %f' % (-np.log(0.1)))

```

```

loss: 2.353468
sanity check: 2.302585

```

1.2 Inline Question 1:

Why do we expect our loss to be close to $-\log(0.1)$? Explain briefly.**

Your answer: Noting that the softmax loss can be thought of as assigning a probability to each class, then because the weight matrix W is randomly initialized, we should expect that it would result in an image being classified as any of the 10 possible classes with equal probability 0.1. The full loss is then minus the log of the average over all images, which should also be approximately 0.1, plus the regularization, which is on the order of $(0.0001)^2 \ll -\log(0.1)$ as a result of the random initialization of the weights on the order of 0.0001.

```
In [4]: # Complete the implementation of softmax_loss_naive and implement a (naive)
# version of the gradient that uses nested loops.
loss, grad = softmax_loss_naive(W, X_dev, y_dev, 0.0)

# As we did for the SVM, use numeric gradient checking as a debugging tool.
# The numeric gradient should be close to the analytic gradient.
from cs682.gradient_check import grad_check_sparse
f = lambda w: softmax_loss_naive(w, X_dev, y_dev, 0.0)[0]
grad_numerical = grad_check_sparse(f, W, grad, 10)

# similar to SVM case, do another gradient check with regularization
loss, grad = softmax_loss_naive(W, X_dev, y_dev, 5e1)
f = lambda w: softmax_loss_naive(w, X_dev, y_dev, 5e1)[0]
grad_numerical = grad_check_sparse(f, W, grad, 10)
```

```
numerical: -3.783523 analytic: -3.783523, relative error: 2.151180e-09
numerical: -2.485062 analytic: -2.485062, relative error: 9.946774e-09
numerical: -4.928226 analytic: -4.928226, relative error: 2.419355e-09
numerical: 0.729395 analytic: 0.729395, relative error: 3.767954e-08
numerical: -1.275667 analytic: -1.275667, relative error: 3.307935e-08
numerical: -3.449171 analytic: -3.449171, relative error: 2.204763e-08
numerical: 1.801794 analytic: 1.801794, relative error: 1.951253e-08
numerical: 2.776163 analytic: 2.776163, relative error: 1.750058e-08
numerical: 2.622774 analytic: 2.622774, relative error: 1.454267e-08
numerical: 0.896589 analytic: 0.896589, relative error: 7.427041e-08
numerical: 0.569886 analytic: 0.569886, relative error: 8.111644e-08
numerical: -0.375095 analytic: -0.375095, relative error: 1.529064e-07
numerical: 2.061371 analytic: 2.061371, relative error: 2.841968e-08
numerical: -0.124883 analytic: -0.124883, relative error: 2.551571e-08
numerical: -3.025586 analytic: -3.025586, relative error: 6.282732e-09
numerical: 3.302104 analytic: 3.302104, relative error: 1.357320e-08
numerical: -2.217739 analytic: -2.217739, relative error: 1.847363e-08
numerical: -0.026059 analytic: -0.026059, relative error: 6.275273e-08
numerical: 0.267677 analytic: 0.267677, relative error: 1.645450e-07
numerical: -0.824773 analytic: -0.824773, relative error: 5.432262e-08
```

```
In [5]: # Now that we have a naive implementation of the softmax loss function and its gradient,
# implement a vectorized version in softmax_loss_vectorized.
# The two versions should compute the same results, but the vectorized version should be
# much faster.
tic = time.time()
loss_naive, grad_naive = softmax_loss_naive(W, X_dev, y_dev, 0.000005)
toc = time.time()
print('naive loss: %e computed in %fs' % (loss_naive, toc - tic))

from cs682.classifiers.softmax import softmax_loss_vectorized
tic = time.time()
loss_vectorized, grad_vectorized = softmax_loss_vectorized(W, X_dev, y_dev, 0.000005)
toc = time.time()
print('vectorized loss: %e computed in %fs' % (loss_vectorized, toc - tic))

# As we did for the SVM, we use the Frobenius norm to compare the two versions
# of the gradient.
grad_difference = np.linalg.norm(grad_naive - grad_vectorized, ord='fro')
print('Loss difference: %f' % np.abs(loss_naive - loss_vectorized))
print('Gradient difference: %f' % grad_difference)
```

```
naive loss: 2.353468e+00 computed in 0.296276s
vectorized loss: 2.353468e+00 computed in 0.011307s
Loss difference: 0.000000
Gradient difference: 0.000000
```

```
In [9]: # Use the validation set to tune hyperparameters (regularization strength and
# learning rate). You should experiment with different ranges for the learning
# rates and regularization strengths; if you are careful you should be able to
# get a classification accuracy of over 0.35 on the validation set.
```

```

from cs682.classifiers import Softmax
results = {}
best_val = -1
best_softmax = None
learning_rates = [5e-8, 1e-7, 2.5e-7, 5e-7, 1e-6, 2.5e-6, 5e-6]
regularization_strengths = [5e3, 1e4, 2.5e4, 5e4, 1e5]

#####
# TODO:
# Use the validation set to set the learning rate and regularization strength. #
# This should be identical to the validation that you did for the SVM; save #
# the best trained softmax classifier in best_softmax. #
#####

for rate in learning_rates:
    for strength in regularization_strengths:
        softmax_val = Softmax()
        loss_hist = softmax_val.train(X_train,
                                      y_train,
                                      learning_rate=rate,
                                      reg=strength,
                                      num_iters=5000,
                                      verbose=False)

        yt_pred = softmax_val.predict(X_train)
        yt_acc = np.mean(y_train == yt_pred)
        yv_pred = softmax_val.predict(X_val)
        yv_acc = np.mean(y_val == yv_pred)

        print('rate = %e, strength = %e, train acc = %f, val acc = %f' % (rate,
                                strength, yt_acc, yv_acc))

        results[(rate, strength)] = (yt_acc, yv_acc)

        if yv_acc > best_val:
            best_val = yv_acc
            best_softmax = softmax_val

#####
# END OF YOUR CODE
#####

# # Print out results.
# for lr, reg in sorted(results):
#     train_accuracy, val_accuracy = results[(lr, reg)]
#     print('lr %e reg %e train accuracy: %f val accuracy: %f' % (
#         lr, reg, train_accuracy, val_accuracy))

print('best validation accuracy achieved during cross-validation: %f' % best_val)

rate = 5.000000e-08, strength = 5.000000e+03, train acc = 0.369367, val acc = 0.375000
rate = 5.000000e-08, strength = 1.000000e+04, train acc = 0.357388, val acc = 0.374000
rate = 5.000000e-08, strength = 2.500000e+04, train acc = 0.329531, val acc = 0.343000
rate = 5.000000e-08, strength = 5.000000e+04, train acc = 0.305612, val acc = 0.323000
rate = 5.000000e-08, strength = 1.000000e+05, train acc = 0.284918, val acc = 0.295000
rate = 1.000000e-07, strength = 5.000000e+03, train acc = 0.373143, val acc = 0.395000
rate = 1.000000e-07, strength = 1.000000e+04, train acc = 0.356653, val acc = 0.374000
rate = 1.000000e-07, strength = 2.500000e+04, train acc = 0.332286, val acc = 0.345000
rate = 1.000000e-07, strength = 5.000000e+04, train acc = 0.303061, val acc = 0.314000
rate = 1.000000e-07, strength = 1.000000e+05, train acc = 0.282633, val acc = 0.290000
rate = 2.500000e-07, strength = 5.000000e+03, train acc = 0.374755, val acc = 0.390000
rate = 2.500000e-07, strength = 1.000000e+04, train acc = 0.354388, val acc = 0.358000
rate = 2.500000e-07, strength = 2.500000e+04, train acc = 0.331020, val acc = 0.345000
rate = 2.500000e-07, strength = 5.000000e+04, train acc = 0.295265, val acc = 0.319000
rate = 2.500000e-07, strength = 1.000000e+05, train acc = 0.290469, val acc = 0.307000
rate = 5.000000e-07, strength = 5.000000e+03, train acc = 0.367918, val acc = 0.382000
rate = 5.000000e-07, strength = 1.000000e+04, train acc = 0.349653, val acc = 0.372000
rate = 5.000000e-07, strength = 2.500000e+04, train acc = 0.323122, val acc = 0.345000
rate = 5.000000e-07, strength = 5.000000e+04, train acc = 0.300163, val acc = 0.306000
rate = 5.000000e-07, strength = 1.000000e+05, train acc = 0.281122, val acc = 0.294000
rate = 1.000000e-06, strength = 5.000000e+03, train acc = 0.359653, val acc = 0.370000
rate = 1.000000e-06, strength = 1.000000e+04, train acc = 0.348122, val acc = 0.359000
rate = 1.000000e-06, strength = 2.500000e+04, train acc = 0.322204, val acc = 0.332000
rate = 1.000000e-06, strength = 5.000000e+04, train acc = 0.297633, val acc = 0.297000
rate = 1.000000e-06, strength = 1.000000e+05, train acc = 0.286449, val acc = 0.300000
rate = 2.500000e-06, strength = 5.000000e+03, train acc = 0.354204, val acc = 0.371000

```



```

rate = 2.500000e-06, strength = 1.000000e+04, train acc = 0.334816, val acc = 0.342000
rate = 2.500000e-06, strength = 2.500000e+04, train acc = 0.299898, val acc = 0.297000
rate = 2.500000e-06, strength = 5.000000e+04, train acc = 0.278796, val acc = 0.299000
rate = 2.500000e-06, strength = 1.000000e+05, train acc = 0.221469, val acc = 0.233000
rate = 5.000000e-06, strength = 5.000000e+03, train acc = 0.303020, val acc = 0.321000
rate = 5.000000e-06, strength = 1.000000e+04, train acc = 0.262694, val acc = 0.279000
rate = 5.000000e-06, strength = 2.500000e+04, train acc = 0.197469, val acc = 0.217000
rate = 5.000000e-06, strength = 5.000000e+04, train acc = 0.155776, val acc = 0.152000
rate = 5.000000e-06, strength = 1.000000e+05, train acc = 0.091143, val acc = 0.095000
best validation accuracy achieved during cross-validation: 0.395000

```

```

In [11]: # evaluate on test set
# Evaluate the best softmax on test set
y_test_pred = best_softmax.predict(X_test)
test_accuracy = np.mean(y_test == y_test_pred)
print('softmax on raw pixels final test set accuracy: %f' % (test_accuracy, ))

```

softmax on raw pixels final test set accuracy: 0.378000

Inline Question - True or False

It's possible to add a new datapoint to a training set that would leave the SVM loss unchanged, but this is not the case with the Softmax classifier loss.

Your answer: False

Your explanation: For SVM loss, a new image could be added to the training set whose score with respect to its correct class is never greater than its scores with respect to all other incorrect classes by the margin specified in the loss function, therefore the SVM loss will be unchanged. However, Softmax loss does not “filter” scores in this way, therefore any new image in the training set will contribute to the total loss.

```

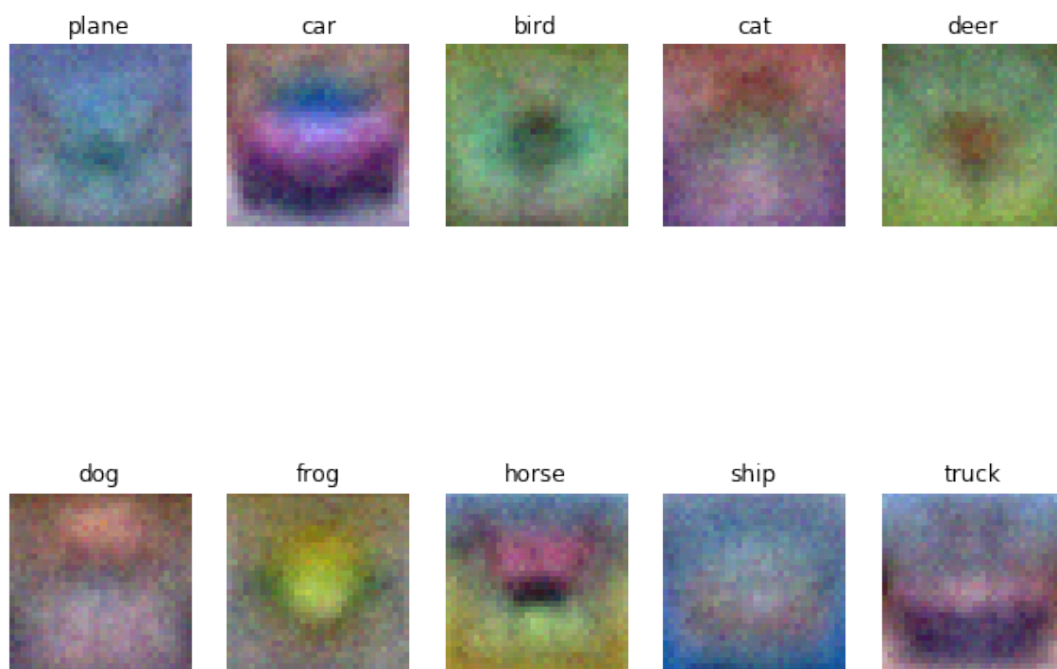
In [12]: # Visualize the learned weights for each class
w = best_softmax.W[:-1,:] # strip out the bias
w = w.reshape(32, 32, 3, 10)

w_min, w_max = np.min(w), np.max(w)

classes = ['plane', 'car', 'bird', 'cat', 'deer', 'dog', 'frog', 'horse', 'ship',
'truck']
for i in range(10):
    plt.subplot(2, 5, i + 1)

    # Rescale the weights to be between 0 and 255
    wimg = 255.0 * (w[:, :, :, i].squeeze() - w_min) / (w_max - w_min)
    plt.imshow(wimg.astype('uint8'))
    plt.axis('off')
    plt.title(classes[i])

```



In []:

two_layer_net

September 24, 2019

1 Implementing a Neural Network

In this exercise we will develop a neural network with fully-connected layers to perform classification, and test it out on the CIFAR-10 dataset.

In [1]: *# A bit of setup*

```
import numpy as np
import matplotlib.pyplot as plt

from cs682.classifiers.neural_net import TwoLayerNet

from __future__ import print_function

%matplotlib inline
plt.rcParams['figure.figsize'] = (10.0, 8.0) # set default size of plots
plt.rcParams['image.interpolation'] = 'nearest'
plt.rcParams['image.cmap'] = 'gray'

# for auto-reloading external modules
# see http://stackoverflow.com/questions/1907993/autoreload-of-modules-in-ipython
%load_ext autoreload
%autoreload 2

def rel_error(x, y):
    """ returns relative error """
    return np.max(np.abs(x - y) / (np.maximum(1e-8, np.abs(x) + np.abs(y))))
```

We will use the class `TwoLayerNet` in the file `cs682/classifiers/neural_net.py` to represent instances of our network. The network parameters are stored in the instance variable `self.params` where keys are string parameter names and values are numpy arrays. Below, we initialize toy data and a toy model that we will use to develop your implementation.

In [2]: *# Create a small net and some toy data to check your implementations.*
Note that we set the random seed for repeatable experiments.

```
input_size = 4
hidden_size = 10
num_classes = 3
num_inputs = 5

def init_toy_model():
    np.random.seed(0)
    return TwoLayerNet(input_size, hidden_size, num_classes, std=1e-1)

def init_toy_data():
    np.random.seed(1)
    X = 10 * np.random.randn(num_inputs, input_size)
    y = np.array([0, 1, 2, 2, 1])
    return X, y

net = init_toy_model()
X, y = init_toy_data()
```

2 Forward pass: compute scores

Open the file `cs682/classifiers/neural_net.py` and look at the method `TwoLayerNet.loss`. This function is very similar to the loss functions you have written for the SVM and Softmax exercises: It takes the

data and weights and computes the class scores, the loss, and the gradients on the parameters.

Implement the first part of the forward pass which uses the weights and biases to compute the scores for all inputs.

```
In [3]: scores = net.loss(X)
        print('Your scores:')
        print(scores)
        print()
        print('correct scores:')
        correct_scores = np.asarray([
            [-0.81233741, -1.27654624, -0.70335995],
            [-0.17129677, -1.18803311, -0.47310444],
            [-0.51590475, -1.01354314, -0.8504215 ],
            [-0.15419291, -0.48629638, -0.52901952],
            [-0.00618733, -0.12435261, -0.15226949]])
        print(correct_scores)
        print()

        # The difference should be very small. We get < 1e-7
        print('Difference between your scores and correct scores:')
        print(np.sum(np.abs(scores - correct_scores)))
```

```
Your scores:
[[-0.81233741 -1.27654624 -0.70335995]
 [-0.17129677 -1.18803311 -0.47310444]
 [-0.51590475 -1.01354314 -0.8504215 ]
 [-0.15419291 -0.48629638 -0.52901952]
 [-0.00618733 -0.12435261 -0.15226949]]
```

```
correct scores:
[[-0.81233741 -1.27654624 -0.70335995]
 [-0.17129677 -1.18803311 -0.47310444]
 [-0.51590475 -1.01354314 -0.8504215 ]
 [-0.15419291 -0.48629638 -0.52901952]
 [-0.00618733 -0.12435261 -0.15226949]]
```

```
Difference between your scores and correct scores:
3.6802720745909845e-08
```

3 Forward pass: compute loss

In the same function, implement the second part that computes the data and regularization loss.

```
In [4]: loss, _ = net.loss(X, y, reg=0.05)
        correct_loss = 1.30378789133
        print('loss = %f' % loss)
        print('correct loss = %f' % correct_loss)

        # should be very small, we get < 1e-12
        print('Difference between your loss and correct loss:')
        print(np.sum(np.abs(loss - correct_loss)))
```

```
loss = 1.303788
correct loss = 1.303788
Difference between your loss and correct loss:
1.794120407794253e-13
```

4 Backward pass

Implement the rest of the function. This will compute the gradient of the loss with respect to the variables $W1$, $b1$, $W2$, and $b2$. Now that you (hopefully!) have a correctly implemented forward pass, you can debug your backward pass using a numeric gradient check:

```
In [5]: from cs682.gradient_check import eval_numerical_gradient

        # Use numeric gradient checking to check your implementation of the backward pass.
        # If your implementation is correct, the difference between the numeric and
        # analytic gradients should be less than 1e-8 for each of  $W1$ ,  $W2$ ,  $b1$ , and  $b2$ .
```

```

loss, grads = net.loss(X, y, reg=0.05)
# print(grads)

# these should all be less than 1e-8 or so
for param_name in grads:
    f = lambda W: net.loss(X, y, reg=0.05)[0]
    param_grad_num = eval_numerical_gradient(f, net.params[param_name], verbose=False)
    # print(grads[param_name])
    # print(param_grad_num)
    # print(param_grad_num - grads[param_name])
    print('%s max relative error: %e' % (param_name, rel_error(param_grad_num,
grads[param_name])))

b2 max relative error: 4.447625e-11
W2 max relative error: 3.440708e-09
W1 max relative error: 3.561318e-09
b1 max relative error: 2.738421e-09

```

5 Train the network

To train the network we will use stochastic gradient descent (SGD), similar to the SVM and Softmax classifiers. Look at the function `TwoLayerNet.train` and fill in the missing sections to implement the training procedure. This should be very similar to the training procedure you used for the SVM and Softmax classifiers. You will also have to implement `TwoLayerNet.predict`, as the training process periodically performs prediction to keep track of accuracy over time while the network trains.

Once you have implemented the method, run the code below to train a two-layer network on toy data. You should achieve a training loss less than 0.2.

```

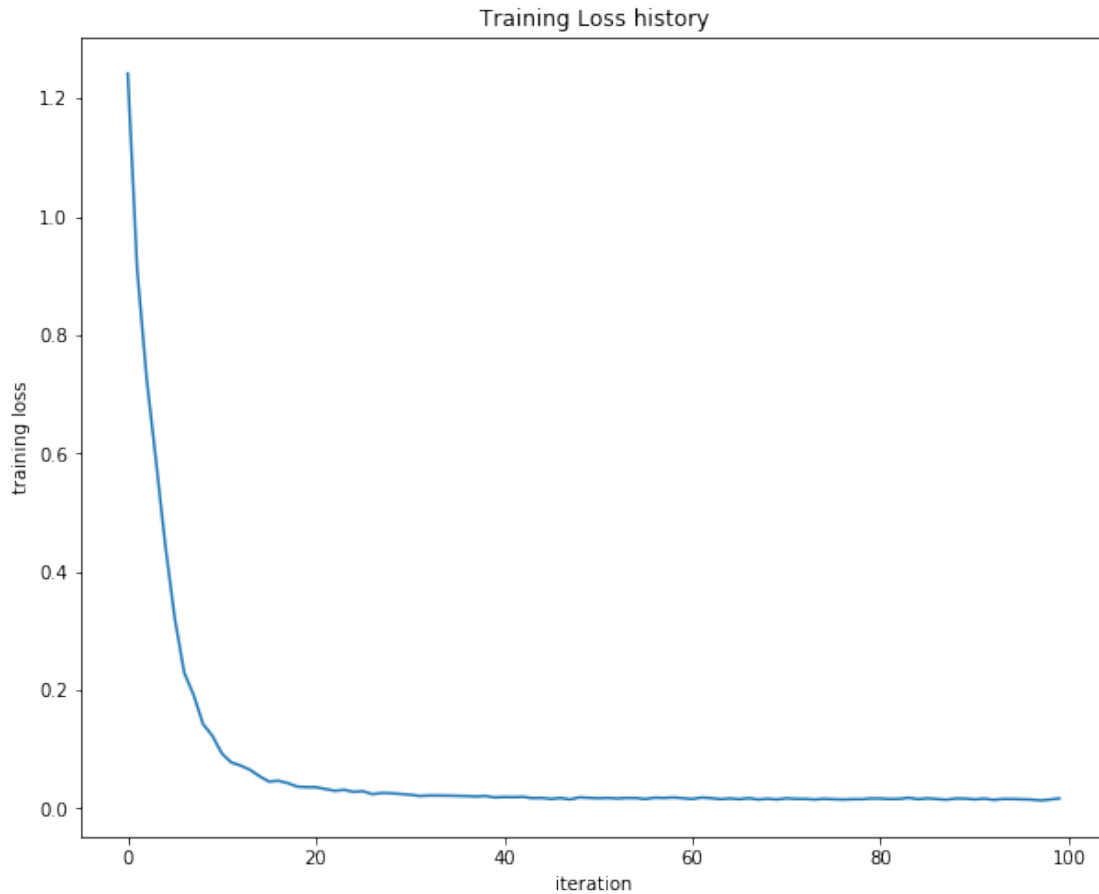
In [6]: net = init_toy_model()
        stats = net.train(X, y, X, y,
                          learning_rate=1e-1, reg=5e-6,
                          num_iters=100, verbose=False)

        print('Final training loss: ', stats['loss_history'][-1])

        # plot the loss history
        plt.plot(stats['loss_history'])
        plt.xlabel('iteration')
        plt.ylabel('training loss')
        plt.title('Training Loss history')
        plt.show()

```

Final training loss: 0.017149607938732093



6 Load the data

Now that you have implemented a two-layer network that passes gradient checks and works on toy data, it's time to load up our favorite CIFAR-10 data so we can use it to train a classifier on a real dataset.

```
In [7]: from cs682.data_utils import load_CIFAR10

def get_CIFAR10_data(num_training=49000, num_validation=1000, num_test=1000):
    """
    Load the CIFAR-10 dataset from disk and perform preprocessing to prepare
    it for the two-layer neural net classifier. These are the same steps as
    we used for the SVM, but condensed to a single function.
    """
    # Load the raw CIFAR-10 data
    cifar10_dir = 'cs682/datasets/cifar-10-batches-py'

    X_train, y_train, X_test, y_test = load_CIFAR10(cifar10_dir)

    # Subsample the data
    mask = list(range(num_training, num_training + num_validation))
    X_val = X_train[mask]
    y_val = y_train[mask]
    mask = list(range(num_training))
    X_train = X_train[mask]
    y_train = y_train[mask]
    mask = list(range(num_test))
    X_test = X_test[mask]
    y_test = y_test[mask]

    # Normalize the data: subtract the mean image
```

```

mean_image = np.mean(X_train, axis=0)
X_train -= mean_image
X_val -= mean_image
X_test -= mean_image

# Reshape data to rows
X_train = X_train.reshape(num_training, -1)
X_val = X_val.reshape(num_validation, -1)
X_test = X_test.reshape(num_test, -1)

return X_train, y_train, X_val, y_val, X_test, y_test

# Cleaning up variables to prevent loading data multiple times (which may cause memory
issue)
try:
    del X_train, y_train
    del X_test, y_test
    print('Clear previously loaded data.')
except:
    pass

# Invoke the above function to get our data.
X_train, y_train, X_val, y_val, X_test, y_test = get_CIFAR10_data()
print('Train data shape: ', X_train.shape)
print('Train labels shape: ', y_train.shape)
print('Validation data shape: ', X_val.shape)
print('Validation labels shape: ', y_val.shape)
print('Test data shape: ', X_test.shape)
print('Test labels shape: ', y_test.shape)

```

```

Train data shape: (49000, 3072)
Train labels shape: (49000,)
Validation data shape: (1000, 3072)
Validation labels shape: (1000,)
Test data shape: (1000, 3072)
Test labels shape: (1000,)

```

7 Train a network

To train our network we will use SGD. In addition, we will adjust the learning rate with an exponential learning rate schedule as optimization proceeds; after each epoch, we will reduce the learning rate by multiplying it by a decay rate.

```

In [8]: input_size = 32 * 32 * 3
        hidden_size = 50
        num_classes = 10
        net = TwoLayerNet(input_size, hidden_size, num_classes)

# Train the network
stats = net.train(X_train, y_train, X_val, y_val,
                  num_iters=1000, batch_size=200,
                  learning_rate=1e-4, learning_rate_decay=0.95,
                  reg=0.25, verbose=True)

# Predict on the validation set
val_acc = (net.predict(X_val) == y_val).mean()
print('Validation accuracy: ', val_acc)

```

```

iteration 0 / 1000: loss 2.302954
iteration 100 / 1000: loss 2.302550
iteration 200 / 1000: loss 2.297648
iteration 300 / 1000: loss 2.259602
iteration 400 / 1000: loss 2.204170
iteration 500 / 1000: loss 2.118565
iteration 600 / 1000: loss 2.051535
iteration 700 / 1000: loss 1.988466
iteration 800 / 1000: loss 2.006591
iteration 900 / 1000: loss 1.951473
Validation accuracy: 0.287

```

8 Debug the training

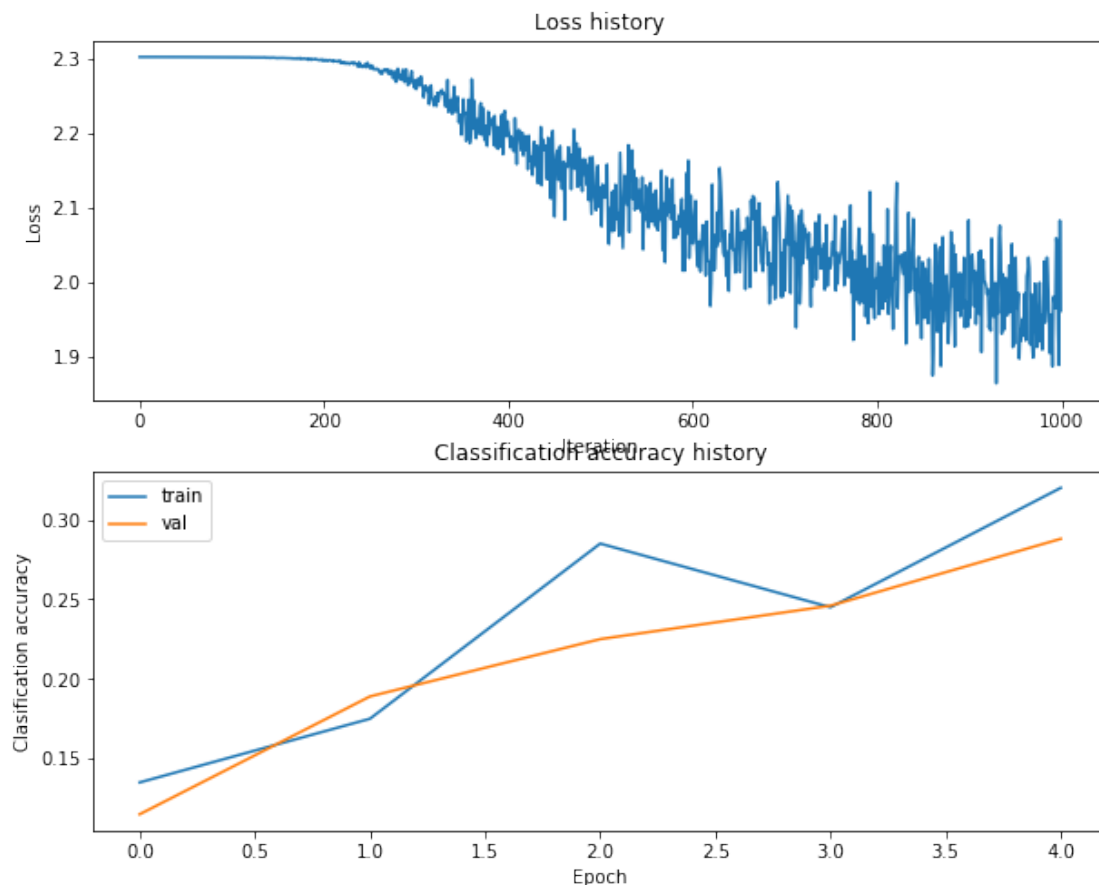
With the default parameters we provided above, you should get a validation accuracy of about 0.29 on the validation set. This isn't very good.

One strategy for getting insight into what's wrong is to plot the loss function and the accuracies on the training and validation sets during optimization.

Another strategy is to visualize the weights that were learned in the first layer of the network. In most neural networks trained on visual data, the first layer weights typically show some visible structure when visualized.

```
In [9]: # Plot the loss function and train / validation accuracies
plt.subplot(2, 1, 1)
plt.plot(stats['loss_history'])
plt.title('Loss history')
plt.xlabel('Iteration')
plt.ylabel('Loss')

plt.subplot(2, 1, 2)
plt.plot(stats['train_acc_history'], label='train')
plt.plot(stats['val_acc_history'], label='val')
plt.title('Classification accuracy history')
plt.xlabel('Epoch')
plt.ylabel('Classification accuracy')
plt.legend()
plt.show()
```



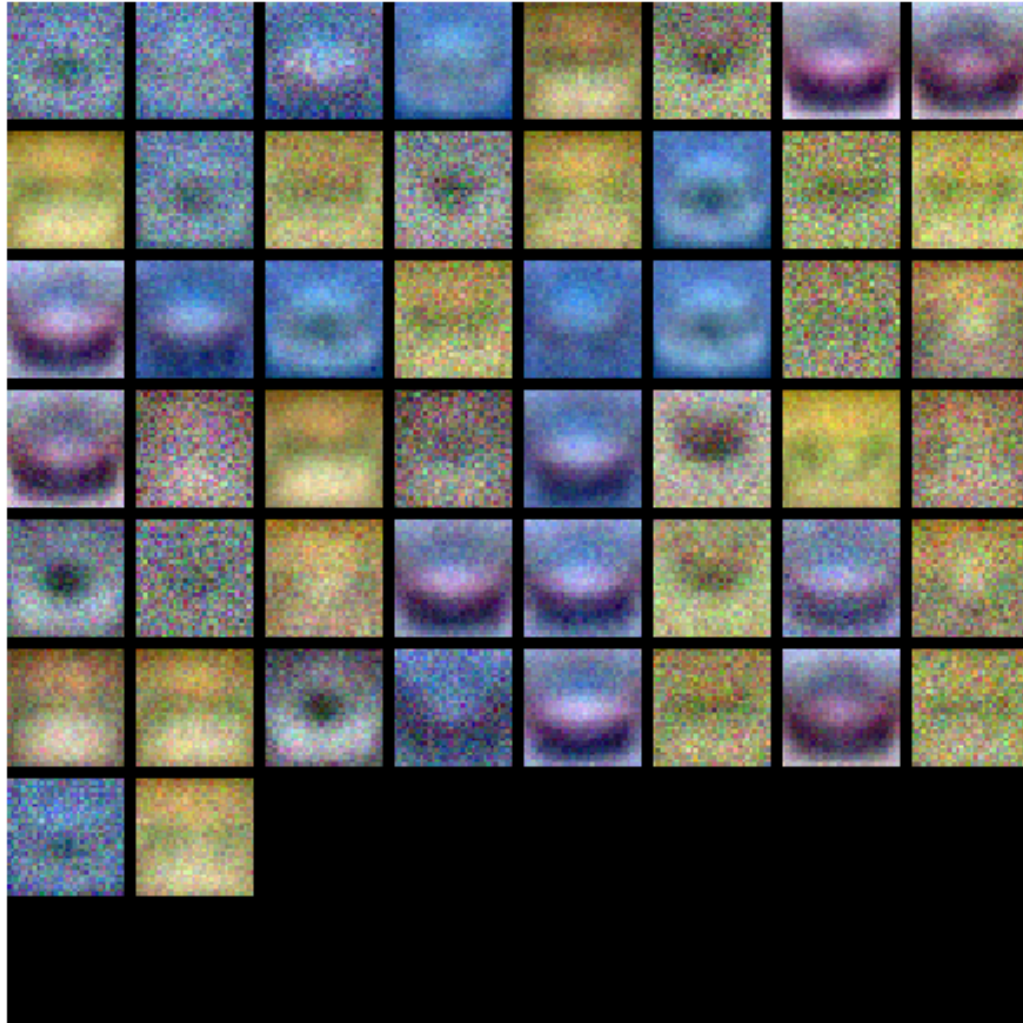
```
In [10]: from cs682.vis_utils import visualize_grid

# Visualize the weights of the network
```



```
def show_net_weights(net):
    W1 = net.params['W1']
    W1 = W1.reshape(32, 32, 3, -1).transpose(3, 0, 1, 2)
    plt.imshow(visualize_grid(W1, padding=3).astype('uint8'))
    plt.gca().axis('off')
    plt.show()

show_net_weights(net)
```



9 Tune your hyperparameters

What's wrong?. Looking at the visualizations above, we see that the loss is decreasing more or less linearly, which seems to suggest that the learning rate may be too low. Moreover, there is no gap between the training and validation accuracy, suggesting that the model we used has low capacity, and that we should increase its size. On the other hand, with a very large model we would expect to see more overfitting, which would manifest itself as a very large gap between the training and validation accuracy.

Tuning. Tuning the hyperparameters and developing intuition for how they affect the final performance is a large part of using Neural Networks, so we want you to get a lot of practice. Below, you should experiment with different values of the various hyperparameters, including hidden layer size, learning rate, number of training epochs, and regularization strength. You might also consider tuning the learning rate decay, but you should be able to get good performance using the default value.

Approximate results. You should be able to aim to achieve a classification accuracy of greater than 48% on the validation set. Our best network gets over 52% on the validation set.

Experiment: Your goal in this exercise is to get as good of a result on CIFAR-10 as you can, with a fully-connected Neural Network. Feel free to implement your own techniques (e.g. PCA to reduce dimensionality, or adding dropout, or adding features to the solver, etc.).

```
In [77]: #####
# TODO: Tune hyperparameters using the validation set. Store your best trained #
# model in best_net.                                                         #
#                                                                            #
# To help debug your network, it may help to use visualizations similar to the #
# ones we used above; these visualizations will have significant qualitative    #
# differences from the ones we saw above for the poorly tuned network.        #
#                                                                            #
# Tweaking hyperparameters by hand can be fun, but you might find it useful to #
# write code to sweep through possible combinations of hyperparameters        #
# automatically like we did on the previous exercises.                       #
#####

def select_hyper_parameters():
    results = {}
    best_acc = -1
    best_net = None

    learning_rates = [2e-3, 2.5e-3, 3e-3]
    regularization_strengths = [1e-5, 2.5e-5, 5e-5, 1e-4, 2.5e-4, 5e-4]
    hidden_layer_sizes = [100, 250, 500, 750]
    batch_sizes = [250, 500]
    num_epochs = 20

    for rate in learning_rates:
        for strength in regularization_strengths:
            for units in hidden_layer_sizes:
                for batch_size in batch_sizes:

                    num_iters = num_epochs * int(X_train.shape[0] / batch_size)

                    net = TwoLayerNet(input_size, units, num_classes)

                    stats = net.train(X_train, y_train, X_val, y_val,
                                     learning_rate=rate,
                                     reg=strength,
                                     batch_size=batch_size,
                                     num_iters=num_iters)

                    val_acc = (net.predict(X_val) == y_val).mean()
                    print('rate: %f, strength: %.7f, hidden units: %d, batch size: %d,
acc: %f' % (rate,
              strength,
              units,
              batch_size,
              val_acc))

                    if val_acc > best_acc:
                        best_acc = val_acc
                        best_net = net

                    results[(rate, strength, units, batch_size, num_iters)] = (val_acc,
stats)

    return best_net, best_acc, results

# the hyperparameters below were selected based on a previous call to the function
# select_hyper_parameters defined above, which will not run by default when this cell is
# evaluated.
#
# rate: 0.003000, strength: 0.0000500, layers: 500, batch size: 500, epochs: 20, acc:
0.563000

# if you wish to perform the full hyperparameter search, change the value of the
variable
```

```

# perform_cross_validation to True and evaluate the cell again. cross validation will
take over
# an hour to complete
perform_cross_validation = False

if perform_cross_validation is True:
    best_net, best_acc, results = select_hyper_parameters()
else:
    # optimal hyperparameters found using cross-validation function defined above
    best_learning_rate = 3e-3
    best_reg_strength = 5e-4
    best_num_hidden_units = 500
    best_batch_size = 500

    best_net = TwoLayerNet(input_size, best_num_hidden_units, num_classes)

    num_iters = 20 * int(X_train.shape[0] / best_batch_size)

    stats = net.train(X_train, y_train, X_val, y_val,
                      learning_rate=best_learning_rate,
                      reg=best_reg_strength,
                      batch_size=best_batch_size,
                      num_iters=num_iters)

    best_acc = (best_net.predict(X_val) == y_val).mean()

    print('rate: %f, strength: %.7f, hidden units: %d, batch size: %d, acc: %f' %
          (best_learning_rate,
           best_reg_strength,
           best_num_hidden_units,
           best_batch_size,
           best_acc))

```

```

#####
#                               END OF YOUR CODE                               #
#####

```

```

rate: 0.002000, strength: 0.0000100, layers: 100, batch size: 250, epochs: 20, acc: 0.503000
rate: 0.002000, strength: 0.0000100, layers: 100, batch size: 500, epochs: 20, acc: 0.507000
rate: 0.002000, strength: 0.0000100, layers: 250, batch size: 250, epochs: 20, acc: 0.529000
rate: 0.002000, strength: 0.0000100, layers: 250, batch size: 500, epochs: 20, acc: 0.541000
rate: 0.002000, strength: 0.0000100, layers: 500, batch size: 250, epochs: 20, acc: 0.545000
rate: 0.002000, strength: 0.0000100, layers: 500, batch size: 500, epochs: 20, acc: 0.553000
rate: 0.002000, strength: 0.0000100, layers: 750, batch size: 250, epochs: 20, acc: 0.542000
rate: 0.002000, strength: 0.0000100, layers: 750, batch size: 500, epochs: 20, acc: 0.529000
rate: 0.002000, strength: 0.0000250, layers: 100, batch size: 250, epochs: 20, acc: 0.501000
rate: 0.002000, strength: 0.0000250, layers: 100, batch size: 500, epochs: 20, acc: 0.520000
rate: 0.002000, strength: 0.0000250, layers: 250, batch size: 250, epochs: 20, acc: 0.542000
rate: 0.002000, strength: 0.0000250, layers: 250, batch size: 500, epochs: 20, acc: 0.552000
rate: 0.002000, strength: 0.0000250, layers: 500, batch size: 250, epochs: 20, acc: 0.529000
rate: 0.002000, strength: 0.0000250, layers: 500, batch size: 500, epochs: 20, acc: 0.547000
rate: 0.002000, strength: 0.0000250, layers: 750, batch size: 250, epochs: 20, acc: 0.533000
rate: 0.002000, strength: 0.0000250, layers: 750, batch size: 500, epochs: 20, acc: 0.528000
rate: 0.002000, strength: 0.0000500, layers: 100, batch size: 250, epochs: 20, acc: 0.538000
rate: 0.002000, strength: 0.0000500, layers: 100, batch size: 500, epochs: 20, acc: 0.517000
rate: 0.002000, strength: 0.0000500, layers: 250, batch size: 250, epochs: 20, acc: 0.535000
rate: 0.002000, strength: 0.0000500, layers: 250, batch size: 500, epochs: 20, acc: 0.545000
rate: 0.002000, strength: 0.0000500, layers: 500, batch size: 250, epochs: 20, acc: 0.512000
rate: 0.002000, strength: 0.0000500, layers: 500, batch size: 500, epochs: 20, acc: 0.536000
rate: 0.002000, strength: 0.0000500, layers: 750, batch size: 250, epochs: 20, acc: 0.532000
rate: 0.002000, strength: 0.0000500, layers: 750, batch size: 500, epochs: 20, acc: 0.536000
rate: 0.002000, strength: 0.0001000, layers: 100, batch size: 250, epochs: 20, acc: 0.505000
rate: 0.002000, strength: 0.0001000, layers: 100, batch size: 500, epochs: 20, acc: 0.526000
rate: 0.002000, strength: 0.0001000, layers: 250, batch size: 250, epochs: 20, acc: 0.538000
rate: 0.002000, strength: 0.0001000, layers: 250, batch size: 500, epochs: 20, acc: 0.548000
rate: 0.002000, strength: 0.0001000, layers: 500, batch size: 250, epochs: 20, acc: 0.546000
rate: 0.002000, strength: 0.0001000, layers: 500, batch size: 500, epochs: 20, acc: 0.542000
rate: 0.002000, strength: 0.0001000, layers: 750, batch size: 250, epochs: 20, acc: 0.525000
rate: 0.002000, strength: 0.0001000, layers: 750, batch size: 500, epochs: 20, acc: 0.534000
rate: 0.002000, strength: 0.0002500, layers: 100, batch size: 250, epochs: 20, acc: 0.515000
rate: 0.002000, strength: 0.0002500, layers: 100, batch size: 500, epochs: 20, acc: 0.531000
rate: 0.002000, strength: 0.0002500, layers: 250, batch size: 250, epochs: 20, acc: 0.540000
rate: 0.002000, strength: 0.0002500, layers: 250, batch size: 500, epochs: 20, acc: 0.547000
rate: 0.002000, strength: 0.0002500, layers: 500, batch size: 250, epochs: 20, acc: 0.529000

```

[illegible]

```

rate: 0.003000, strength: 0.0000250, layers: 100, batch size: 500, epochs: 20, acc: 0.533000
rate: 0.003000, strength: 0.0000250, layers: 250, batch size: 250, epochs: 20, acc: 0.512000
rate: 0.003000, strength: 0.0000250, layers: 250, batch size: 500, epochs: 20, acc: 0.548000
rate: 0.003000, strength: 0.0000250, layers: 500, batch size: 250, epochs: 20, acc: 0.544000
rate: 0.003000, strength: 0.0000250, layers: 500, batch size: 500, epochs: 20, acc: 0.557000
rate: 0.003000, strength: 0.0000250, layers: 750, batch size: 250, epochs: 20, acc: 0.517000
rate: 0.003000, strength: 0.0000250, layers: 750, batch size: 500, epochs: 20, acc: 0.541000
rate: 0.003000, strength: 0.0000500, layers: 100, batch size: 250, epochs: 20, acc: 0.518000
rate: 0.003000, strength: 0.0000500, layers: 100, batch size: 500, epochs: 20, acc: 0.533000
rate: 0.003000, strength: 0.0000500, layers: 250, batch size: 250, epochs: 20, acc: 0.543000
rate: 0.003000, strength: 0.0000500, layers: 250, batch size: 500, epochs: 20, acc: 0.528000
rate: 0.003000, strength: 0.0000500, layers: 500, batch size: 250, epochs: 20, acc: 0.539000
rate: 0.003000, strength: 0.0000500, layers: 500, batch size: 500, epochs: 20, acc: 0.563000
rate: 0.003000, strength: 0.0000500, layers: 750, batch size: 250, epochs: 20, acc: 0.533000
rate: 0.003000, strength: 0.0000500, layers: 750, batch size: 500, epochs: 20, acc: 0.555000
rate: 0.003000, strength: 0.0001000, layers: 100, batch size: 250, epochs: 20, acc: 0.531000
rate: 0.003000, strength: 0.0001000, layers: 100, batch size: 500, epochs: 20, acc: 0.504000
rate: 0.003000, strength: 0.0001000, layers: 250, batch size: 250, epochs: 20, acc: 0.512000
rate: 0.003000, strength: 0.0001000, layers: 250, batch size: 500, epochs: 20, acc: 0.530000
rate: 0.003000, strength: 0.0001000, layers: 500, batch size: 250, epochs: 20, acc: 0.526000
rate: 0.003000, strength: 0.0001000, layers: 500, batch size: 500, epochs: 20, acc: 0.543000
rate: 0.003000, strength: 0.0001000, layers: 750, batch size: 250, epochs: 20, acc: 0.549000
rate: 0.003000, strength: 0.0001000, layers: 750, batch size: 500, epochs: 20, acc: 0.510000
rate: 0.003000, strength: 0.0002500, layers: 100, batch size: 250, epochs: 20, acc: 0.512000
rate: 0.003000, strength: 0.0002500, layers: 100, batch size: 500, epochs: 20, acc: 0.503000
rate: 0.003000, strength: 0.0002500, layers: 250, batch size: 250, epochs: 20, acc: 0.515000
rate: 0.003000, strength: 0.0002500, layers: 250, batch size: 500, epochs: 20, acc: 0.519000
rate: 0.003000, strength: 0.0002500, layers: 500, batch size: 250, epochs: 20, acc: 0.483000
rate: 0.003000, strength: 0.0002500, layers: 500, batch size: 500, epochs: 20, acc: 0.545000
rate: 0.003000, strength: 0.0002500, layers: 750, batch size: 250, epochs: 20, acc: 0.525000
rate: 0.003000, strength: 0.0002500, layers: 750, batch size: 500, epochs: 20, acc: 0.563000
rate: 0.003000, strength: 0.0005000, layers: 100, batch size: 250, epochs: 20, acc: 0.513000
rate: 0.003000, strength: 0.0005000, layers: 100, batch size: 500, epochs: 20, acc: 0.507000
rate: 0.003000, strength: 0.0005000, layers: 250, batch size: 250, epochs: 20, acc: 0.528000
rate: 0.003000, strength: 0.0005000, layers: 250, batch size: 500, epochs: 20, acc: 0.513000
rate: 0.003000, strength: 0.0005000, layers: 500, batch size: 250, epochs: 20, acc: 0.529000
rate: 0.003000, strength: 0.0005000, layers: 500, batch size: 500, epochs: 20, acc: 0.541000
rate: 0.003000, strength: 0.0005000, layers: 750, batch size: 250, epochs: 20, acc: 0.538000
rate: 0.003000, strength: 0.0005000, layers: 750, batch size: 500, epochs: 20, acc: 0.544000

```

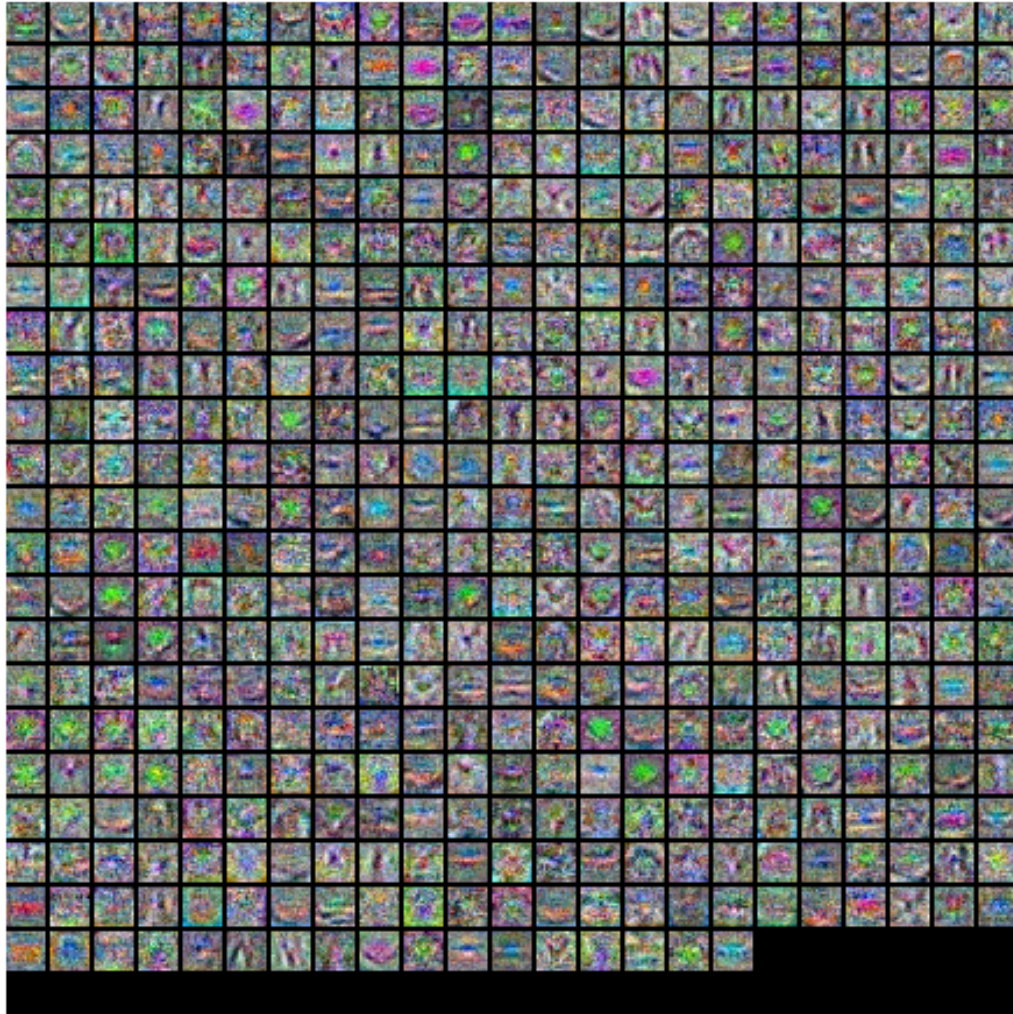
```

In [84]: # visualize the weights of the best network
print(best_acc)
show_net_weights(best_net)

# rate: 0.003000, strength: 0.0000500, layers: 500, batch size: 500, epochs: 20, acc:
0.563000

```

0.563



10 Run on the test set

When you are done experimenting, you should evaluate your final trained network on the test set; you should get above 48%.

```
In [85]: test_acc = (best_net.predict(X_test) == y_test).mean()
         print('Test accuracy: ', test_acc)
```

Test accuracy: 0.531

Inline Question

Now that you have trained a Neural Network classifier, you may find that your testing accuracy is much lower than the training accuracy. In what ways can we decrease this gap? Select all that apply. 1. Train on a larger dataset. 2. Add more hidden units. 3. Increase the regularization strength. 4. None of the above.

Your answer: 1, 3

Your explanation: High training accuracy and low test accuracy suggests high variance and over-fitting to the data, i.e. that the neural network is learning the training data's features as well as it's noise, causing it to generalize poorly to unseen data.

By increasing the size of the training dataset, we are increasing the number of instances from which the network can learn the features of the data, while at the same time decreasing the variance of the data, i.e. minimizing the amount that can be learned from the noise.

If our network is already overfitting to the data, increasing the number of hidden units is likely to only make the issue worse; by increasing the amount of nonlinearity in the hidden layer, we are essentially allowing the neural network to learn the noise of the data more thoroughly, resulting in an even lower testing accuracy.

By increasing the regularization strength, we are incentivizing the neural network to learn a weight matrix with many small and zero entries, rather than one with many large and non-zero entries. The former is more likely to generalize well to unseen data because it represents a simpler model of the features in the training data.

In []:

features

September 24, 2019

1 Image features exercise

Complete and hand in this [completed worksheet](#) (including its outputs and any supporting code outside of the worksheet) with your assignment submission. For more details see the [assignments page](#) on the course website.

We have seen that we can achieve reasonable performance on an image classification task by training a linear classifier on the pixels of the input image. In this exercise we will show that we can improve our classification performance by training linear classifiers not on raw pixels but on features that are computed from the raw pixels.

All of your work for this exercise will be done in this notebook.

```
In [1]: import random
import numpy as np
from cs682.data_utils import load_CIFAR10
import matplotlib.pyplot as plt

from __future__ import print_function

%matplotlib inline
plt.rcParams['figure.figsize'] = (10.0, 8.0) # set default size of plots
plt.rcParams['image.interpolation'] = 'nearest'
plt.rcParams['image.cmap'] = 'gray'

# for auto-reloading external modules
# see http://stackoverflow.com/questions/1907993/autoreload-of-modules-in-ipython
%load_ext autoreload
%autoreload 2
```

1.1 Load data

Similar to previous exercises, we will load CIFAR-10 data from disk.

```
In [2]: from cs682.features import color_histogram_hsv, hog_feature

def get_CIFAR10_data(num_training=49000, num_validation=1000, num_test=1000):
    # Load the raw CIFAR-10 data
    cifar10_dir = 'cs682/datasets/cifar-10-batches-py'

    X_train, y_train, X_test, y_test = load_CIFAR10(cifar10_dir)

    # Subsample the data
    mask = list(range(num_training, num_training + num_validation))
    X_val = X_train[mask]
    y_val = y_train[mask]
    mask = list(range(num_training))
    X_train = X_train[mask]
    y_train = y_train[mask]
    mask = list(range(num_test))
    X_test = X_test[mask]
    y_test = y_test[mask]

    return X_train, y_train, X_val, y_val, X_test, y_test

# Cleaning up variables to prevent loading data multiple times (which may cause memory
# issue)
try:
    del X_train, y_train
```



```

del X_test, y_test
print('Clear previously loaded data.')
except:
    pass

X_train, y_train, X_val, y_val, X_test, y_test = get_CIFAR10_data()

```

1.2 Extract Features

For each image we will compute a Histogram of Oriented Gradients (HOG) as well as a color histogram using the hue channel in HSV color space. We form our final feature vector for each image by concatenating the HOG and color histogram feature vectors.

Roughly speaking, HOG should capture the texture of the image while ignoring color information, and the color histogram represents the color of the input image while ignoring texture. As a result, we expect that using both together ought to work better than using either alone. Verifying this assumption would be a good thing to try for your interests.

The `hog_feature` and `color_histogram_hsv` functions both operate on a single image and return a feature vector for that image. The `extract_features` function takes a set of images and a list of feature functions and evaluates each feature function on each image, storing the results in a matrix where each column is the concatenation of all feature vectors for a single image.

```

In [6]: from cs682.features import *

num_color_bins = 10 # Number of bins in the color histogram
feature_fns = [hog_feature, lambda img: color_histogram_hsv(img, nbin=num_color_bins)]
X_train_feats = extract_features(X_train, feature_fns, verbose=True)
X_val_feats = extract_features(X_val, feature_fns)
X_test_feats = extract_features(X_test, feature_fns)

# Preprocessing: Subtract the mean feature
mean_feat = np.mean(X_train_feats, axis=0, keepdims=True)
X_train_feats -= mean_feat
X_val_feats -= mean_feat
X_test_feats -= mean_feat

# Preprocessing: Divide by standard deviation. This ensures that each feature
# has roughly the same scale.
std_feat = np.std(X_train_feats, axis=0, keepdims=True)
X_train_feats /= std_feat
X_val_feats /= std_feat
X_test_feats /= std_feat

# Preprocessing: Add a bias dimension
X_train_feats = np.hstack([X_train_feats, np.ones((X_train_feats.shape[0], 1))])
X_val_feats = np.hstack([X_val_feats, np.ones((X_val_feats.shape[0], 1))])
X_test_feats = np.hstack([X_test_feats, np.ones((X_test_feats.shape[0], 1))])

```

```

Done extracting features for 1000 / 49000 images
Done extracting features for 2000 / 49000 images
Done extracting features for 3000 / 49000 images
Done extracting features for 4000 / 49000 images
Done extracting features for 5000 / 49000 images
Done extracting features for 6000 / 49000 images
Done extracting features for 7000 / 49000 images
Done extracting features for 8000 / 49000 images
Done extracting features for 9000 / 49000 images
Done extracting features for 10000 / 49000 images
Done extracting features for 11000 / 49000 images
Done extracting features for 12000 / 49000 images
Done extracting features for 13000 / 49000 images
Done extracting features for 14000 / 49000 images
Done extracting features for 15000 / 49000 images
Done extracting features for 16000 / 49000 images
Done extracting features for 17000 / 49000 images
Done extracting features for 18000 / 49000 images
Done extracting features for 19000 / 49000 images
Done extracting features for 20000 / 49000 images
Done extracting features for 21000 / 49000 images
Done extracting features for 22000 / 49000 images

```

```

Done extracting features for 23000 / 49000 images
Done extracting features for 24000 / 49000 images
Done extracting features for 25000 / 49000 images
Done extracting features for 26000 / 49000 images
Done extracting features for 27000 / 49000 images
Done extracting features for 28000 / 49000 images
Done extracting features for 29000 / 49000 images
Done extracting features for 30000 / 49000 images
Done extracting features for 31000 / 49000 images
Done extracting features for 32000 / 49000 images
Done extracting features for 33000 / 49000 images
Done extracting features for 34000 / 49000 images
Done extracting features for 35000 / 49000 images
Done extracting features for 36000 / 49000 images
Done extracting features for 37000 / 49000 images
Done extracting features for 38000 / 49000 images
Done extracting features for 39000 / 49000 images
Done extracting features for 40000 / 49000 images
Done extracting features for 41000 / 49000 images
Done extracting features for 42000 / 49000 images
Done extracting features for 43000 / 49000 images
Done extracting features for 44000 / 49000 images
Done extracting features for 45000 / 49000 images
Done extracting features for 46000 / 49000 images
Done extracting features for 47000 / 49000 images
Done extracting features for 48000 / 49000 images

```

1.3 Train SVM on features

Using the multiclass SVM code developed earlier in the assignment, train SVMs on top of the features extracted above; this should achieve better results than training SVMs directly on top of raw pixels.

```

In [15]: # Use the validation set to tune the learning rate and regularization strength

from cs682.classifiers.linear_classifier import LinearSVM

f1 = lambda x: 1*(10**x)
f2 = lambda x: 2.5*(10**x)
f3 = lambda x: 5*(10**x)

learning_rates = [f(e) for e in [-7, -6, -5, -4, -3] for f in (f1, f2, f3)]
regularization_strengths = [f(e) for e in [4, 5, 6] for f in (f1, f2, f3)]
batch_sizes = [250, 500, 750]
num_iterations = [250, 500, 1000, 2000]

results = {}
best_val = -1
best_svm = None

#####
# TODO:
# Use the validation set to set the learning rate and regularization strength. #
# This should be identical to the validation that you did for the SVM; save #
# the best trained classifier in best_svm. You might also want to play #
# with different numbers of bins in the color histogram. If you are careful #
# you should be able to get accuracy of near 0.44 on the validation set. #
#####

for rate in learning_rates:
    for strength in regularization_strengths:
        for batch_size in batch_sizes:
            for num_iters in num_iterations:

                svm = LinearSVM()
                loss_history = svm.train(X_train_feats,
                                        y_train,
                                        learning_rate=rate,
                                        reg=strength,
                                        num_iters=num_iters,
                                        batch_size=batch_size)

                yt_pred = svm.predict(X_train_feats)
                yt_acc = np.mean(y_train == yt_pred)
                yv_pred = svm.predict(X_val_feats)

```

```

        yv_acc = np.mean(y_val == yv_pred)

        results[(rate, strength, batch_size, num_iters)] = (yt_acc, yv_acc)

        if yv_acc > best_val:
            best_val = yv_acc
            best_svm = svm

#####
#                               END OF YOUR CODE                               #
#####

# Print out results.
for lr, reg, batch, num_iter in sorted(results):
    train_accuracy, val_accuracy = results[(lr, reg, batch, num_iter)]
    print('lr %e reg %e batch %d iters %d train accuracy: %f val accuracy: %f' % (
        lr, reg, batch, num_iter, train_accuracy, val_accuracy))

print('best validation accuracy achieved during cross-validation: %f' % best_val)

/home/brendan/Desktop/school/grad/f19/cs682/assignment1/cs682/classifiers/linear_svm.py:144: RuntimeWarning: overflow
encountered in multiply
    dW = X.T.dot(M) / num_train + 2 * reg * W
/home/brendan/Desktop/school/grad/f19/cs682/assignment1/cs682/classifiers/linear_svm.py:110: RuntimeWarning: invalid
value encountered in subtract
    losses = all_scores - correct_class_scores + 1
/home/brendan/Desktop/school/grad/f19/cs682/assignment1/cs682/classifiers/linear_classifier.py:70: RuntimeWarning:
invalid value encountered in subtract
    self.W -= learning_rate * grad
/home/brendan/Desktop/school/grad/f19/cs682/assignment1/cs682/classifiers/linear_svm.py:116: RuntimeWarning: invalid
value encountered in greater
    margins = np.where(losses > 0, losses, 0)
/home/brendan/Desktop/school/grad/f19/cs682/assignment1/env/lib/python3.5/site-packages/numpy/core/fromnumeric.py:90:
RuntimeWarning: overflow encountered in reduce
    return ufunc.reduce(obj, axis, dtype, out, **passkwargs)

lr 1.000000e-07 reg 1.000000e+01 batch 250 iters 250 train accuracy: 0.106327 val accuracy: 0.102000
lr 1.000000e-07 reg 1.000000e+01 batch 250 iters 500 train accuracy: 0.123592 val accuracy: 0.117000
lr 1.000000e-07 reg 1.000000e+01 batch 250 iters 1000 train accuracy: 0.128510 val accuracy: 0.138000
lr 1.000000e-07 reg 1.000000e+01 batch 250 iters 2000 train accuracy: 0.091980 val accuracy: 0.102000
lr 1.000000e-07 reg 1.000000e+01 batch 500 iters 250 train accuracy: 0.124061 val accuracy: 0.143000
lr 1.000000e-07 reg 1.000000e+01 batch 500 iters 500 train accuracy: 0.116633 val accuracy: 0.121000
lr 1.000000e-07 reg 1.000000e+01 batch 500 iters 1000 train accuracy: 0.113265 val accuracy: 0.094000
lr 1.000000e-07 reg 1.000000e+01 batch 500 iters 2000 train accuracy: 0.126163 val accuracy: 0.137000
lr 1.000000e-07 reg 1.000000e+01 batch 750 iters 250 train accuracy: 0.076837 val accuracy: 0.088000
lr 1.000000e-07 reg 1.000000e+01 batch 750 iters 500 train accuracy: 0.101551 val accuracy: 0.126000
lr 1.000000e-07 reg 1.000000e+01 batch 750 iters 1000 train accuracy: 0.089755 val accuracy: 0.128000
lr 1.000000e-07 reg 1.000000e+01 batch 750 iters 2000 train accuracy: 0.125857 val accuracy: 0.118000
lr 1.000000e-07 reg 2.500000e+01 batch 250 iters 250 train accuracy: 0.081551 val accuracy: 0.074000
lr 1.000000e-07 reg 2.500000e+01 batch 250 iters 500 train accuracy: 0.113735 val accuracy: 0.099000
lr 1.000000e-07 reg 2.500000e+01 batch 250 iters 1000 train accuracy: 0.111551 val accuracy: 0.117000
lr 1.000000e-07 reg 2.500000e+01 batch 250 iters 2000 train accuracy: 0.131286 val accuracy: 0.128000
lr 1.000000e-07 reg 2.500000e+01 batch 500 iters 250 train accuracy: 0.099306 val accuracy: 0.103000
lr 1.000000e-07 reg 2.500000e+01 batch 500 iters 500 train accuracy: 0.104939 val accuracy: 0.112000
lr 1.000000e-07 reg 2.500000e+01 batch 500 iters 1000 train accuracy: 0.098327 val accuracy: 0.114000
lr 1.000000e-07 reg 2.500000e+01 batch 500 iters 2000 train accuracy: 0.123735 val accuracy: 0.126000
lr 1.000000e-07 reg 2.500000e+01 batch 750 iters 250 train accuracy: 0.088224 val accuracy: 0.077000
lr 1.000000e-07 reg 2.500000e+01 batch 750 iters 500 train accuracy: 0.092612 val accuracy: 0.090000
lr 1.000000e-07 reg 2.500000e+01 batch 750 iters 1000 train accuracy: 0.119020 val accuracy: 0.147000
lr 1.000000e-07 reg 2.500000e+01 batch 750 iters 2000 train accuracy: 0.116776 val accuracy: 0.104000
lr 1.000000e-07 reg 5.000000e+01 batch 250 iters 250 train accuracy: 0.109367 val accuracy: 0.103000
lr 1.000000e-07 reg 5.000000e+01 batch 250 iters 500 train accuracy: 0.106245 val accuracy: 0.098000
lr 1.000000e-07 reg 5.000000e+01 batch 250 iters 1000 train accuracy: 0.109510 val accuracy: 0.104000
lr 1.000000e-07 reg 5.000000e+01 batch 250 iters 2000 train accuracy: 0.119082 val accuracy: 0.121000
lr 1.000000e-07 reg 5.000000e+01 batch 500 iters 250 train accuracy: 0.096184 val accuracy: 0.099000
lr 1.000000e-07 reg 5.000000e+01 batch 500 iters 500 train accuracy: 0.110327 val accuracy: 0.114000
lr 1.000000e-07 reg 5.000000e+01 batch 500 iters 1000 train accuracy: 0.111531 val accuracy: 0.127000
lr 1.000000e-07 reg 5.000000e+01 batch 500 iters 2000 train accuracy: 0.117163 val accuracy: 0.130000
lr 1.000000e-07 reg 5.000000e+01 batch 750 iters 250 train accuracy: 0.119000 val accuracy: 0.114000
lr 1.000000e-07 reg 5.000000e+01 batch 750 iters 500 train accuracy: 0.111510 val accuracy: 0.116000
lr 1.000000e-07 reg 5.000000e+01 batch 750 iters 1000 train accuracy: 0.115735 val accuracy: 0.106000

```


lr 1.000000e-07 reg 5.000000e+03 batch 500 iters 2000 train accuracy: 0.195224 val accuracy: 0.187000
 lr 1.000000e-07 reg 5.000000e+03 batch 750 iters 250 train accuracy: 0.125286 val accuracy: 0.115000
 lr 1.000000e-07 reg 5.000000e+03 batch 750 iters 500 train accuracy: 0.104918 val accuracy: 0.099000
 lr 1.000000e-07 reg 5.000000e+03 batch 750 iters 1000 train accuracy: 0.114959 val accuracy: 0.099000
 lr 1.000000e-07 reg 5.000000e+03 batch 750 iters 2000 train accuracy: 0.171224 val accuracy: 0.172000
 lr 1.000000e-07 reg 1.000000e+04 batch 250 iters 250 train accuracy: 0.109000 val accuracy: 0.132000
 lr 1.000000e-07 reg 1.000000e+04 batch 250 iters 500 train accuracy: 0.111816 val accuracy: 0.120000
 lr 1.000000e-07 reg 1.000000e+04 batch 250 iters 1000 train accuracy: 0.120184 val accuracy: 0.125000
 lr 1.000000e-07 reg 1.000000e+04 batch 250 iters 2000 train accuracy: 0.310469 val accuracy: 0.295000
 lr 1.000000e-07 reg 1.000000e+04 batch 500 iters 250 train accuracy: 0.103980 val accuracy: 0.101000
 lr 1.000000e-07 reg 1.000000e+04 batch 500 iters 500 train accuracy: 0.105286 val accuracy: 0.112000
 lr 1.000000e-07 reg 1.000000e+04 batch 500 iters 1000 train accuracy: 0.158857 val accuracy: 0.156000
 lr 1.000000e-07 reg 1.000000e+04 batch 500 iters 2000 train accuracy: 0.318469 val accuracy: 0.327000
 lr 1.000000e-07 reg 1.000000e+04 batch 750 iters 250 train accuracy: 0.111020 val accuracy: 0.102000
 lr 1.000000e-07 reg 1.000000e+04 batch 750 iters 500 train accuracy: 0.119959 val accuracy: 0.117000
 lr 1.000000e-07 reg 1.000000e+04 batch 750 iters 1000 train accuracy: 0.118204 val accuracy: 0.115000
 lr 1.000000e-07 reg 1.000000e+04 batch 750 iters 2000 train accuracy: 0.333061 val accuracy: 0.340000
 lr 1.000000e-07 reg 2.500000e+04 batch 250 iters 250 train accuracy: 0.099735 val accuracy: 0.126000
 lr 1.000000e-07 reg 2.500000e+04 batch 250 iters 500 train accuracy: 0.138571 val accuracy: 0.137000
 lr 1.000000e-07 reg 2.500000e+04 batch 250 iters 1000 train accuracy: 0.343469 val accuracy: 0.350000
 lr 1.000000e-07 reg 2.500000e+04 batch 250 iters 2000 train accuracy: 0.414265 val accuracy: 0.415000
 lr 1.000000e-07 reg 2.500000e+04 batch 500 iters 250 train accuracy: 0.110633 val accuracy: 0.104000
 lr 1.000000e-07 reg 2.500000e+04 batch 500 iters 500 train accuracy: 0.136531 val accuracy: 0.145000
 lr 1.000000e-07 reg 2.500000e+04 batch 500 iters 1000 train accuracy: 0.335959 val accuracy: 0.327000
 lr 1.000000e-07 reg 2.500000e+04 batch 500 iters 2000 train accuracy: 0.415816 val accuracy: 0.419000
 lr 1.000000e-07 reg 2.500000e+04 batch 750 iters 250 train accuracy: 0.107612 val accuracy: 0.133000
 lr 1.000000e-07 reg 2.500000e+04 batch 750 iters 500 train accuracy: 0.130204 val accuracy: 0.121000
 lr 1.000000e-07 reg 2.500000e+04 batch 750 iters 1000 train accuracy: 0.346429 val accuracy: 0.327000
 lr 1.000000e-07 reg 2.500000e+04 batch 750 iters 2000 train accuracy: 0.414347 val accuracy: 0.420000
 lr 1.000000e-07 reg 5.000000e+04 batch 250 iters 250 train accuracy: 0.094735 val accuracy: 0.095000
 lr 1.000000e-07 reg 5.000000e+04 batch 250 iters 500 train accuracy: 0.234837 val accuracy: 0.223000
 lr 1.000000e-07 reg 5.000000e+04 batch 250 iters 1000 train accuracy: 0.415878 val accuracy: 0.416000
 lr 1.000000e-07 reg 5.000000e+04 batch 250 iters 2000 train accuracy: 0.412000 val accuracy: 0.409000
 lr 1.000000e-07 reg 5.000000e+04 batch 500 iters 250 train accuracy: 0.122612 val accuracy: 0.110000
 lr 1.000000e-07 reg 5.000000e+04 batch 500 iters 500 train accuracy: 0.240959 val accuracy: 0.242000
 lr 1.000000e-07 reg 5.000000e+04 batch 500 iters 1000 train accuracy: 0.414449 val accuracy: 0.409000
 lr 1.000000e-07 reg 5.000000e+04 batch 500 iters 2000 train accuracy: 0.414163 val accuracy: 0.424000
 lr 1.000000e-07 reg 5.000000e+04 batch 750 iters 250 train accuracy: 0.096918 val accuracy: 0.096000
 lr 1.000000e-07 reg 5.000000e+04 batch 750 iters 500 train accuracy: 0.268245 val accuracy: 0.274000
 lr 1.000000e-07 reg 5.000000e+04 batch 750 iters 1000 train accuracy: 0.414837 val accuracy: 0.418000
 lr 1.000000e-07 reg 5.000000e+04 batch 750 iters 2000 train accuracy: 0.415041 val accuracy: 0.415000
 lr 1.000000e-07 reg 1.000000e+05 batch 250 iters 250 train accuracy: 0.209000 val accuracy: 0.221000
 lr 1.000000e-07 reg 1.000000e+05 batch 250 iters 500 train accuracy: 0.414714 val accuracy: 0.422000
 lr 1.000000e-07 reg 1.000000e+05 batch 250 iters 1000 train accuracy: 0.411469 val accuracy: 0.418000
 lr 1.000000e-07 reg 1.000000e+05 batch 250 iters 2000 train accuracy: 0.418367 val accuracy: 0.424000
 lr 1.000000e-07 reg 1.000000e+05 batch 500 iters 250 train accuracy: 0.205490 val accuracy: 0.218000
 lr 1.000000e-07 reg 1.000000e+05 batch 500 iters 500 train accuracy: 0.411980 val accuracy: 0.417000
 lr 1.000000e-07 reg 1.000000e+05 batch 500 iters 1000 train accuracy: 0.415531 val accuracy: 0.426000
 lr 1.000000e-07 reg 1.000000e+05 batch 500 iters 2000 train accuracy: 0.416224 val accuracy: 0.418000
 lr 1.000000e-07 reg 1.000000e+05 batch 750 iters 250 train accuracy: 0.190755 val accuracy: 0.197000
 lr 1.000000e-07 reg 1.000000e+05 batch 750 iters 500 train accuracy: 0.415531 val accuracy: 0.421000
 lr 1.000000e-07 reg 1.000000e+05 batch 750 iters 1000 train accuracy: 0.417714 val accuracy: 0.424000
 lr 1.000000e-07 reg 1.000000e+05 batch 750 iters 2000 train accuracy: 0.412510 val accuracy: 0.415000
 lr 1.000000e-07 reg 2.500000e+05 batch 250 iters 250 train accuracy: 0.416224 val accuracy: 0.416000
 lr 1.000000e-07 reg 2.500000e+05 batch 250 iters 500 train accuracy: 0.413735 val accuracy: 0.406000
 lr 1.000000e-07 reg 2.500000e+05 batch 250 iters 1000 train accuracy: 0.417327 val accuracy: 0.430000
 lr 1.000000e-07 reg 2.500000e+05 batch 250 iters 2000 train accuracy: 0.402735 val accuracy: 0.387000
 lr 1.000000e-07 reg 2.500000e+05 batch 500 iters 250 train accuracy: 0.409714 val accuracy: 0.410000
 lr 1.000000e-07 reg 2.500000e+05 batch 500 iters 500 train accuracy: 0.414449 val accuracy: 0.410000
 lr 1.000000e-07 reg 2.500000e+05 batch 500 iters 1000 train accuracy: 0.416184 val accuracy: 0.413000
 lr 1.000000e-07 reg 2.500000e+05 batch 500 iters 2000 train accuracy: 0.413531 val accuracy: 0.417000
 lr 1.000000e-07 reg 2.500000e+05 batch 750 iters 250 train accuracy: 0.415714 val accuracy: 0.419000
 lr 1.000000e-07 reg 2.500000e+05 batch 750 iters 500 train accuracy: 0.416163 val accuracy: 0.415000
 lr 1.000000e-07 reg 2.500000e+05 batch 750 iters 1000 train accuracy: 0.409796 val accuracy: 0.413000
 lr 1.000000e-07 reg 2.500000e+05 batch 750 iters 2000 train accuracy: 0.412245 val accuracy: 0.413000
 lr 1.000000e-07 reg 5.000000e+05 batch 250 iters 250 train accuracy: 0.406735 val accuracy: 0.397000
 lr 1.000000e-07 reg 5.000000e+05 batch 250 iters 500 train accuracy: 0.407265 val accuracy: 0.403000
 lr 1.000000e-07 reg 5.000000e+05 batch 250 iters 1000 train accuracy: 0.397224 val accuracy: 0.405000

lr 1.000000e-07 reg 5.000000e+05 batch 250 iters 2000 train accuracy: 0.404265 val accuracy: 0.400000
 lr 1.000000e-07 reg 5.000000e+05 batch 500 iters 250 train accuracy: 0.406469 val accuracy: 0.392000
 lr 1.000000e-07 reg 5.000000e+05 batch 500 iters 500 train accuracy: 0.407959 val accuracy: 0.409000
 lr 1.000000e-07 reg 5.000000e+05 batch 500 iters 1000 train accuracy: 0.409612 val accuracy: 0.414000
 lr 1.000000e-07 reg 5.000000e+05 batch 500 iters 2000 train accuracy: 0.412551 val accuracy: 0.410000
 lr 1.000000e-07 reg 5.000000e+05 batch 750 iters 250 train accuracy: 0.407265 val accuracy: 0.408000
 lr 1.000000e-07 reg 5.000000e+05 batch 750 iters 500 train accuracy: 0.413347 val accuracy: 0.422000
 lr 1.000000e-07 reg 5.000000e+05 batch 750 iters 1000 train accuracy: 0.413939 val accuracy: 0.402000
 lr 1.000000e-07 reg 5.000000e+05 batch 750 iters 2000 train accuracy: 0.413755 val accuracy: 0.411000
 lr 1.000000e-07 reg 1.000000e+06 batch 250 iters 250 train accuracy: 0.411122 val accuracy: 0.403000
 lr 1.000000e-07 reg 1.000000e+06 batch 250 iters 500 train accuracy: 0.393980 val accuracy: 0.403000
 lr 1.000000e-07 reg 1.000000e+06 batch 250 iters 1000 train accuracy: 0.393592 val accuracy: 0.400000
 lr 1.000000e-07 reg 1.000000e+06 batch 250 iters 2000 train accuracy: 0.408490 val accuracy: 0.408000
 lr 1.000000e-07 reg 1.000000e+06 batch 500 iters 250 train accuracy: 0.411796 val accuracy: 0.402000
 lr 1.000000e-07 reg 1.000000e+06 batch 500 iters 500 train accuracy: 0.406408 val accuracy: 0.403000
 lr 1.000000e-07 reg 1.000000e+06 batch 500 iters 1000 train accuracy: 0.408429 val accuracy: 0.422000
 lr 1.000000e-07 reg 1.000000e+06 batch 500 iters 2000 train accuracy: 0.409878 val accuracy: 0.407000
 lr 1.000000e-07 reg 1.000000e+06 batch 750 iters 250 train accuracy: 0.409796 val accuracy: 0.406000
 lr 1.000000e-07 reg 1.000000e+06 batch 750 iters 500 train accuracy: 0.404796 val accuracy: 0.411000
 lr 1.000000e-07 reg 1.000000e+06 batch 750 iters 1000 train accuracy: 0.412959 val accuracy: 0.411000
 lr 1.000000e-07 reg 1.000000e+06 batch 750 iters 2000 train accuracy: 0.415633 val accuracy: 0.430000
 lr 1.000000e-07 reg 2.500000e+06 batch 250 iters 250 train accuracy: 0.365490 val accuracy: 0.340000
 lr 1.000000e-07 reg 2.500000e+06 batch 250 iters 500 train accuracy: 0.403184 val accuracy: 0.394000
 lr 1.000000e-07 reg 2.500000e+06 batch 250 iters 1000 train accuracy: 0.365816 val accuracy: 0.367000
 lr 1.000000e-07 reg 2.500000e+06 batch 250 iters 2000 train accuracy: 0.354673 val accuracy: 0.353000
 lr 1.000000e-07 reg 2.500000e+06 batch 500 iters 250 train accuracy: 0.386918 val accuracy: 0.383000
 lr 1.000000e-07 reg 2.500000e+06 batch 500 iters 500 train accuracy: 0.392449 val accuracy: 0.371000
 lr 1.000000e-07 reg 2.500000e+06 batch 500 iters 1000 train accuracy: 0.391959 val accuracy: 0.380000
 lr 1.000000e-07 reg 2.500000e+06 batch 500 iters 2000 train accuracy: 0.408551 val accuracy: 0.412000
 lr 1.000000e-07 reg 2.500000e+06 batch 750 iters 250 train accuracy: 0.396449 val accuracy: 0.388000
 lr 1.000000e-07 reg 2.500000e+06 batch 750 iters 500 train accuracy: 0.405735 val accuracy: 0.410000
 lr 1.000000e-07 reg 2.500000e+06 batch 750 iters 1000 train accuracy: 0.398265 val accuracy: 0.379000
 lr 1.000000e-07 reg 2.500000e+06 batch 750 iters 2000 train accuracy: 0.404367 val accuracy: 0.398000
 lr 1.000000e-07 reg 5.000000e+06 batch 250 iters 250 train accuracy: 0.315143 val accuracy: 0.303000
 lr 1.000000e-07 reg 5.000000e+06 batch 250 iters 500 train accuracy: 0.317122 val accuracy: 0.305000
 lr 1.000000e-07 reg 5.000000e+06 batch 250 iters 1000 train accuracy: 0.315816 val accuracy: 0.315000
 lr 1.000000e-07 reg 5.000000e+06 batch 250 iters 2000 train accuracy: 0.342265 val accuracy: 0.346000
 lr 1.000000e-07 reg 5.000000e+06 batch 500 iters 250 train accuracy: 0.356000 val accuracy: 0.376000
 lr 1.000000e-07 reg 5.000000e+06 batch 500 iters 500 train accuracy: 0.367592 val accuracy: 0.343000
 lr 1.000000e-07 reg 5.000000e+06 batch 500 iters 1000 train accuracy: 0.360408 val accuracy: 0.388000
 lr 1.000000e-07 reg 5.000000e+06 batch 500 iters 2000 train accuracy: 0.355796 val accuracy: 0.356000
 lr 1.000000e-07 reg 5.000000e+06 batch 750 iters 250 train accuracy: 0.382469 val accuracy: 0.378000
 lr 1.000000e-07 reg 5.000000e+06 batch 750 iters 500 train accuracy: 0.383020 val accuracy: 0.387000
 lr 1.000000e-07 reg 5.000000e+06 batch 750 iters 1000 train accuracy: 0.372388 val accuracy: 0.371000
 lr 1.000000e-07 reg 5.000000e+06 batch 750 iters 2000 train accuracy: 0.398306 val accuracy: 0.404000
 lr 2.500000e-07 reg 1.000000e+01 batch 250 iters 250 train accuracy: 0.121163 val accuracy: 0.114000
 lr 2.500000e-07 reg 1.000000e+01 batch 250 iters 500 train accuracy: 0.100388 val accuracy: 0.114000
 lr 2.500000e-07 reg 1.000000e+01 batch 250 iters 1000 train accuracy: 0.165367 val accuracy: 0.169000
 lr 2.500000e-07 reg 1.000000e+01 batch 250 iters 2000 train accuracy: 0.178429 val accuracy: 0.173000
 lr 2.500000e-07 reg 1.000000e+01 batch 500 iters 250 train accuracy: 0.097776 val accuracy: 0.098000
 lr 2.500000e-07 reg 1.000000e+01 batch 500 iters 500 train accuracy: 0.118306 val accuracy: 0.109000
 lr 2.500000e-07 reg 1.000000e+01 batch 500 iters 1000 train accuracy: 0.137510 val accuracy: 0.127000
 lr 2.500000e-07 reg 1.000000e+01 batch 500 iters 2000 train accuracy: 0.176694 val accuracy: 0.179000
 lr 2.500000e-07 reg 1.000000e+01 batch 750 iters 250 train accuracy: 0.106551 val accuracy: 0.119000
 lr 2.500000e-07 reg 1.000000e+01 batch 750 iters 500 train accuracy: 0.122224 val accuracy: 0.113000
 lr 2.500000e-07 reg 1.000000e+01 batch 750 iters 1000 train accuracy: 0.134449 val accuracy: 0.134000
 lr 2.500000e-07 reg 1.000000e+01 batch 750 iters 2000 train accuracy: 0.171571 val accuracy: 0.187000
 lr 2.500000e-07 reg 2.500000e+01 batch 250 iters 250 train accuracy: 0.098429 val accuracy: 0.092000
 lr 2.500000e-07 reg 2.500000e+01 batch 250 iters 500 train accuracy: 0.115429 val accuracy: 0.119000
 lr 2.500000e-07 reg 2.500000e+01 batch 250 iters 1000 train accuracy: 0.124837 val accuracy: 0.116000
 lr 2.500000e-07 reg 2.500000e+01 batch 250 iters 2000 train accuracy: 0.159388 val accuracy: 0.167000
 lr 2.500000e-07 reg 2.500000e+01 batch 500 iters 250 train accuracy: 0.114837 val accuracy: 0.119000
 lr 2.500000e-07 reg 2.500000e+01 batch 500 iters 500 train accuracy: 0.119592 val accuracy: 0.116000
 lr 2.500000e-07 reg 2.500000e+01 batch 500 iters 1000 train accuracy: 0.124041 val accuracy: 0.127000
 lr 2.500000e-07 reg 2.500000e+01 batch 500 iters 2000 train accuracy: 0.165469 val accuracy: 0.154000
 lr 2.500000e-07 reg 2.500000e+01 batch 750 iters 250 train accuracy: 0.090980 val accuracy: 0.093000
 lr 2.500000e-07 reg 2.500000e+01 batch 750 iters 500 train accuracy: 0.108429 val accuracy: 0.117000
 lr 2.500000e-07 reg 2.500000e+01 batch 750 iters 1000 train accuracy: 0.123245 val accuracy: 0.144000

lr 2.500000e-07 reg 2.500000e+01 batch 750 iters 2000 train accuracy: 0.151347 val accuracy: 0.145000
 lr 2.500000e-07 reg 5.000000e+01 batch 250 iters 250 train accuracy: 0.084306 val accuracy: 0.096000
 lr 2.500000e-07 reg 5.000000e+01 batch 250 iters 500 train accuracy: 0.134551 val accuracy: 0.133000
 lr 2.500000e-07 reg 5.000000e+01 batch 250 iters 1000 train accuracy: 0.134735 val accuracy: 0.134000
 lr 2.500000e-07 reg 5.000000e+01 batch 250 iters 2000 train accuracy: 0.148633 val accuracy: 0.159000
 lr 2.500000e-07 reg 5.000000e+01 batch 500 iters 250 train accuracy: 0.102857 val accuracy: 0.101000
 lr 2.500000e-07 reg 5.000000e+01 batch 500 iters 500 train accuracy: 0.096510 val accuracy: 0.104000
 lr 2.500000e-07 reg 5.000000e+01 batch 500 iters 1000 train accuracy: 0.150531 val accuracy: 0.129000
 lr 2.500000e-07 reg 5.000000e+01 batch 500 iters 2000 train accuracy: 0.170980 val accuracy: 0.188000
 lr 2.500000e-07 reg 5.000000e+01 batch 750 iters 250 train accuracy: 0.119898 val accuracy: 0.109000
 lr 2.500000e-07 reg 5.000000e+01 batch 750 iters 500 train accuracy: 0.131000 val accuracy: 0.132000
 lr 2.500000e-07 reg 5.000000e+01 batch 750 iters 1000 train accuracy: 0.107980 val accuracy: 0.115000
 lr 2.500000e-07 reg 5.000000e+01 batch 750 iters 2000 train accuracy: 0.138082 val accuracy: 0.151000
 lr 2.500000e-07 reg 1.000000e+02 batch 250 iters 250 train accuracy: 0.085673 val accuracy: 0.073000
 lr 2.500000e-07 reg 1.000000e+02 batch 250 iters 500 train accuracy: 0.142286 val accuracy: 0.136000
 lr 2.500000e-07 reg 1.000000e+02 batch 250 iters 1000 train accuracy: 0.129327 val accuracy: 0.136000
 lr 2.500000e-07 reg 1.000000e+02 batch 250 iters 2000 train accuracy: 0.165041 val accuracy: 0.163000
 lr 2.500000e-07 reg 1.000000e+02 batch 500 iters 250 train accuracy: 0.102000 val accuracy: 0.108000
 lr 2.500000e-07 reg 1.000000e+02 batch 500 iters 500 train accuracy: 0.131592 val accuracy: 0.144000
 lr 2.500000e-07 reg 1.000000e+02 batch 500 iters 1000 train accuracy: 0.115388 val accuracy: 0.116000
 lr 2.500000e-07 reg 1.000000e+02 batch 500 iters 2000 train accuracy: 0.181449 val accuracy: 0.175000
 lr 2.500000e-07 reg 1.000000e+02 batch 750 iters 250 train accuracy: 0.104000 val accuracy: 0.095000
 lr 2.500000e-07 reg 1.000000e+02 batch 750 iters 500 train accuracy: 0.110796 val accuracy: 0.108000
 lr 2.500000e-07 reg 1.000000e+02 batch 750 iters 1000 train accuracy: 0.132816 val accuracy: 0.119000
 lr 2.500000e-07 reg 1.000000e+02 batch 750 iters 2000 train accuracy: 0.170714 val accuracy: 0.171000
 lr 2.500000e-07 reg 2.500000e+02 batch 250 iters 250 train accuracy: 0.100816 val accuracy: 0.107000
 lr 2.500000e-07 reg 2.500000e+02 batch 250 iters 500 train accuracy: 0.095857 val accuracy: 0.104000
 lr 2.500000e-07 reg 2.500000e+02 batch 250 iters 1000 train accuracy: 0.130327 val accuracy: 0.120000
 lr 2.500000e-07 reg 2.500000e+02 batch 250 iters 2000 train accuracy: 0.154571 val accuracy: 0.151000
 lr 2.500000e-07 reg 2.500000e+02 batch 500 iters 250 train accuracy: 0.112082 val accuracy: 0.111000
 lr 2.500000e-07 reg 2.500000e+02 batch 500 iters 500 train accuracy: 0.129918 val accuracy: 0.130000
 lr 2.500000e-07 reg 2.500000e+02 batch 500 iters 1000 train accuracy: 0.104224 val accuracy: 0.110000
 lr 2.500000e-07 reg 2.500000e+02 batch 500 iters 2000 train accuracy: 0.172898 val accuracy: 0.194000
 lr 2.500000e-07 reg 2.500000e+02 batch 750 iters 250 train accuracy: 0.088653 val accuracy: 0.087000
 lr 2.500000e-07 reg 2.500000e+02 batch 750 iters 500 train accuracy: 0.097429 val accuracy: 0.092000
 lr 2.500000e-07 reg 2.500000e+02 batch 750 iters 1000 train accuracy: 0.108939 val accuracy: 0.115000
 lr 2.500000e-07 reg 2.500000e+02 batch 750 iters 2000 train accuracy: 0.165122 val accuracy: 0.172000
 lr 2.500000e-07 reg 5.000000e+02 batch 250 iters 250 train accuracy: 0.112673 val accuracy: 0.097000
 lr 2.500000e-07 reg 5.000000e+02 batch 250 iters 500 train accuracy: 0.109755 val accuracy: 0.112000
 lr 2.500000e-07 reg 5.000000e+02 batch 250 iters 1000 train accuracy: 0.096878 val accuracy: 0.087000
 lr 2.500000e-07 reg 5.000000e+02 batch 250 iters 2000 train accuracy: 0.162571 val accuracy: 0.154000
 lr 2.500000e-07 reg 5.000000e+02 batch 500 iters 250 train accuracy: 0.118143 val accuracy: 0.110000
 lr 2.500000e-07 reg 5.000000e+02 batch 500 iters 500 train accuracy: 0.120776 val accuracy: 0.112000
 lr 2.500000e-07 reg 5.000000e+02 batch 500 iters 1000 train accuracy: 0.146776 val accuracy: 0.142000
 lr 2.500000e-07 reg 5.000000e+02 batch 500 iters 2000 train accuracy: 0.174898 val accuracy: 0.167000
 lr 2.500000e-07 reg 5.000000e+02 batch 750 iters 250 train accuracy: 0.123673 val accuracy: 0.119000
 lr 2.500000e-07 reg 5.000000e+02 batch 750 iters 500 train accuracy: 0.129898 val accuracy: 0.125000
 lr 2.500000e-07 reg 5.000000e+02 batch 750 iters 1000 train accuracy: 0.127265 val accuracy: 0.143000
 lr 2.500000e-07 reg 5.000000e+02 batch 750 iters 2000 train accuracy: 0.186490 val accuracy: 0.193000
 lr 2.500000e-07 reg 1.000000e+03 batch 250 iters 250 train accuracy: 0.117408 val accuracy: 0.119000
 lr 2.500000e-07 reg 1.000000e+03 batch 250 iters 500 train accuracy: 0.130286 val accuracy: 0.130000
 lr 2.500000e-07 reg 1.000000e+03 batch 250 iters 1000 train accuracy: 0.143327 val accuracy: 0.156000
 lr 2.500000e-07 reg 1.000000e+03 batch 250 iters 2000 train accuracy: 0.197224 val accuracy: 0.177000
 lr 2.500000e-07 reg 1.000000e+03 batch 500 iters 250 train accuracy: 0.091612 val accuracy: 0.083000
 lr 2.500000e-07 reg 1.000000e+03 batch 500 iters 500 train accuracy: 0.109959 val accuracy: 0.103000
 lr 2.500000e-07 reg 1.000000e+03 batch 500 iters 1000 train accuracy: 0.143245 val accuracy: 0.127000
 lr 2.500000e-07 reg 1.000000e+03 batch 500 iters 2000 train accuracy: 0.184816 val accuracy: 0.195000
 lr 2.500000e-07 reg 1.000000e+03 batch 750 iters 250 train accuracy: 0.120020 val accuracy: 0.120000
 lr 2.500000e-07 reg 1.000000e+03 batch 750 iters 500 train accuracy: 0.120633 val accuracy: 0.106000
 lr 2.500000e-07 reg 1.000000e+03 batch 750 iters 1000 train accuracy: 0.147796 val accuracy: 0.159000
 lr 2.500000e-07 reg 1.000000e+03 batch 750 iters 2000 train accuracy: 0.214510 val accuracy: 0.209000
 lr 2.500000e-07 reg 2.500000e+03 batch 250 iters 250 train accuracy: 0.095041 val accuracy: 0.095000
 lr 2.500000e-07 reg 2.500000e+03 batch 250 iters 500 train accuracy: 0.106816 val accuracy: 0.102000
 lr 2.500000e-07 reg 2.500000e+03 batch 250 iters 1000 train accuracy: 0.145347 val accuracy: 0.153000
 lr 2.500000e-07 reg 2.500000e+03 batch 250 iters 2000 train accuracy: 0.303429 val accuracy: 0.297000
 lr 2.500000e-07 reg 2.500000e+03 batch 500 iters 250 train accuracy: 0.095082 val accuracy: 0.103000
 lr 2.500000e-07 reg 2.500000e+03 batch 500 iters 500 train accuracy: 0.123408 val accuracy: 0.120000
 lr 2.500000e-07 reg 2.500000e+03 batch 500 iters 1000 train accuracy: 0.137878 val accuracy: 0.131000

lr 2.500000e-07	reg 2.500000e+03	batch 500	iters 2000	train accuracy: 0.310796	val accuracy: 0.310000
lr 2.500000e-07	reg 2.500000e+03	batch 750	iters 250	train accuracy: 0.094347	val accuracy: 0.092000
lr 2.500000e-07	reg 2.500000e+03	batch 750	iters 500	train accuracy: 0.129837	val accuracy: 0.129000
lr 2.500000e-07	reg 2.500000e+03	batch 750	iters 1000	train accuracy: 0.158041	val accuracy: 0.142000
lr 2.500000e-07	reg 2.500000e+03	batch 750	iters 2000	train accuracy: 0.302429	val accuracy: 0.321000
lr 2.500000e-07	reg 5.000000e+03	batch 250	iters 250	train accuracy: 0.112490	val accuracy: 0.108000
lr 2.500000e-07	reg 5.000000e+03	batch 250	iters 500	train accuracy: 0.130265	val accuracy: 0.147000
lr 2.500000e-07	reg 5.000000e+03	batch 250	iters 1000	train accuracy: 0.230184	val accuracy: 0.230000
lr 2.500000e-07	reg 5.000000e+03	batch 250	iters 2000	train accuracy: 0.404857	val accuracy: 0.403000
lr 2.500000e-07	reg 5.000000e+03	batch 500	iters 250	train accuracy: 0.102612	val accuracy: 0.108000
lr 2.500000e-07	reg 5.000000e+03	batch 500	iters 500	train accuracy: 0.142735	val accuracy: 0.143000
lr 2.500000e-07	reg 5.000000e+03	batch 500	iters 1000	train accuracy: 0.235306	val accuracy: 0.266000
lr 2.500000e-07	reg 5.000000e+03	batch 500	iters 2000	train accuracy: 0.414245	val accuracy: 0.414000
lr 2.500000e-07	reg 5.000000e+03	batch 750	iters 250	train accuracy: 0.111673	val accuracy: 0.120000
lr 2.500000e-07	reg 5.000000e+03	batch 750	iters 500	train accuracy: 0.096980	val accuracy: 0.105000
lr 2.500000e-07	reg 5.000000e+03	batch 750	iters 1000	train accuracy: 0.232265	val accuracy: 0.221000
lr 2.500000e-07	reg 5.000000e+03	batch 750	iters 2000	train accuracy: 0.412388	val accuracy: 0.403000
lr 2.500000e-07	reg 1.000000e+04	batch 250	iters 250	train accuracy: 0.112592	val accuracy: 0.097000
lr 2.500000e-07	reg 1.000000e+04	batch 250	iters 500	train accuracy: 0.179429	val accuracy: 0.184000
lr 2.500000e-07	reg 1.000000e+04	batch 250	iters 1000	train accuracy: 0.396449	val accuracy: 0.394000
lr 2.500000e-07	reg 1.000000e+04	batch 250	iters 2000	train accuracy: 0.413980	val accuracy: 0.417000
lr 2.500000e-07	reg 1.000000e+04	batch 500	iters 250	train accuracy: 0.081245	val accuracy: 0.074000
lr 2.500000e-07	reg 1.000000e+04	batch 500	iters 500	train accuracy: 0.176673	val accuracy: 0.180000
lr 2.500000e-07	reg 1.000000e+04	batch 500	iters 1000	train accuracy: 0.396735	val accuracy: 0.397000
lr 2.500000e-07	reg 1.000000e+04	batch 500	iters 2000	train accuracy: 0.416551	val accuracy: 0.417000
lr 2.500000e-07	reg 1.000000e+04	batch 750	iters 250	train accuracy: 0.113694	val accuracy: 0.117000
lr 2.500000e-07	reg 1.000000e+04	batch 750	iters 500	train accuracy: 0.172000	val accuracy: 0.181000
lr 2.500000e-07	reg 1.000000e+04	batch 750	iters 1000	train accuracy: 0.388653	val accuracy: 0.387000
lr 2.500000e-07	reg 1.000000e+04	batch 750	iters 2000	train accuracy: 0.416163	val accuracy: 0.415000
lr 2.500000e-07	reg 2.500000e+04	batch 250	iters 250	train accuracy: 0.145265	val accuracy: 0.173000
lr 2.500000e-07	reg 2.500000e+04	batch 250	iters 500	train accuracy: 0.398939	val accuracy: 0.393000
lr 2.500000e-07	reg 2.500000e+04	batch 250	iters 1000	train accuracy: 0.418041	val accuracy: 0.420000
lr 2.500000e-07	reg 2.500000e+04	batch 250	iters 2000	train accuracy: 0.413816	val accuracy: 0.421000
lr 2.500000e-07	reg 2.500000e+04	batch 500	iters 250	train accuracy: 0.147082	val accuracy: 0.147000
lr 2.500000e-07	reg 2.500000e+04	batch 500	iters 500	train accuracy: 0.402714	val accuracy: 0.404000
lr 2.500000e-07	reg 2.500000e+04	batch 500	iters 1000	train accuracy: 0.416918	val accuracy: 0.423000
lr 2.500000e-07	reg 2.500000e+04	batch 500	iters 2000	train accuracy: 0.413959	val accuracy: 0.418000
lr 2.500000e-07	reg 2.500000e+04	batch 750	iters 250	train accuracy: 0.130653	val accuracy: 0.142000
lr 2.500000e-07	reg 2.500000e+04	batch 750	iters 500	train accuracy: 0.403837	val accuracy: 0.397000
lr 2.500000e-07	reg 2.500000e+04	batch 750	iters 1000	train accuracy: 0.415143	val accuracy: 0.425000
lr 2.500000e-07	reg 2.500000e+04	batch 750	iters 2000	train accuracy: 0.417959	val accuracy: 0.427000
lr 2.500000e-07	reg 5.000000e+04	batch 250	iters 250	train accuracy: 0.378408	val accuracy: 0.379000

lr 2.500000e-07 reg 2.500000e+05 batch 250 iters 2000 train accuracy: 0.407612 val accuracy: 0.406000
 lr 2.500000e-07 reg 2.500000e+05 batch 500 iters 250 train accuracy: 0.412184 val accuracy: 0.403000
 lr 2.500000e-07 reg 2.500000e+05 batch 500 iters 500 train accuracy: 0.408061 val accuracy: 0.401000
 lr 2.500000e-07 reg 2.500000e+05 batch 500 iters 1000 train accuracy: 0.405367 val accuracy: 0.413000
 lr 2.500000e-07 reg 2.500000e+05 batch 500 iters 2000 train accuracy: 0.411388 val accuracy: 0.416000
 lr 2.500000e-07 reg 2.500000e+05 batch 750 iters 250 train accuracy: 0.412429 val accuracy: 0.413000
 lr 2.500000e-07 reg 2.500000e+05 batch 750 iters 500 train accuracy: 0.415000 val accuracy: 0.405000
 lr 2.500000e-07 reg 2.500000e+05 batch 750 iters 1000 train accuracy: 0.413367 val accuracy: 0.413000
 lr 2.500000e-07 reg 2.500000e+05 batch 750 iters 2000 train accuracy: 0.412449 val accuracy: 0.422000
 lr 2.500000e-07 reg 5.000000e+05 batch 250 iters 250 train accuracy: 0.405449 val accuracy: 0.424000
 lr 2.500000e-07 reg 5.000000e+05 batch 250 iters 500 train accuracy: 0.393776 val accuracy: 0.421000
 lr 2.500000e-07 reg 5.000000e+05 batch 250 iters 1000 train accuracy: 0.392510 val accuracy: 0.393000
 lr 2.500000e-07 reg 5.000000e+05 batch 250 iters 2000 train accuracy: 0.387367 val accuracy: 0.370000
 lr 2.500000e-07 reg 5.000000e+05 batch 500 iters 250 train accuracy: 0.405490 val accuracy: 0.417000
 lr 2.500000e-07 reg 5.000000e+05 batch 500 iters 500 train accuracy: 0.408184 val accuracy: 0.418000
 lr 2.500000e-07 reg 5.000000e+05 batch 500 iters 1000 train accuracy: 0.397571 val accuracy: 0.379000
 lr 2.500000e-07 reg 5.000000e+05 batch 500 iters 2000 train accuracy: 0.410245 val accuracy: 0.417000
 lr 2.500000e-07 reg 5.000000e+05 batch 750 iters 250 train accuracy: 0.411000 val accuracy: 0.402000
 lr 2.500000e-07 reg 5.000000e+05 batch 750 iters 500 train accuracy: 0.419408 val accuracy: 0.415000
 lr 2.500000e-07 reg 5.000000e+05 batch 750 iters 1000 train accuracy: 0.406878 val accuracy: 0.392000
 lr 2.500000e-07 reg 5.000000e+05 batch 750 iters 2000 train accuracy: 0.410000 val accuracy: 0.409000
 lr 2.500000e-07 reg 1.000000e+06 batch 250 iters 250 train accuracy: 0.368551 val accuracy: 0.355000
 lr 2.500000e-07 reg 1.000000e+06 batch 250 iters 500 train accuracy: 0.379388 val accuracy: 0.379000
 lr 2.500000e-07 reg 1.000000e+06 batch 250 iters 1000 train accuracy: 0.367082 val accuracy: 0.344000
 lr 2.500000e-07 reg 1.000000e+06 batch 250 iters 2000 train accuracy: 0.372082 val accuracy: 0.379000
 lr 2.500000e-07 reg 1.000000e+06 batch 500 iters 250 train accuracy: 0.406694 val accuracy: 0.409000
 lr 2.500000e-07 reg 1.000000e+06 batch 500 iters 500 train accuracy: 0.395245 val accuracy: 0.391000
 lr 2.500000e-07 reg 1.000000e+06 batch 500 iters 1000 train accuracy: 0.412429 val accuracy: 0.415000
 lr 2.500000e-07 reg 1.000000e+06 batch 500 iters 2000 train accuracy: 0.402224 val accuracy: 0.392000
 lr 2.500000e-07 reg 1.000000e+06 batch 750 iters 250 train accuracy: 0.408878 val accuracy: 0.426000
 lr 2.500000e-07 reg 1.000000e+06 batch 750 iters 500 train accuracy: 0.399224 val accuracy: 0.384000
 lr 2.500000e-07 reg 1.000000e+06 batch 750 iters 1000 train accuracy: 0.398490 val accuracy: 0.387000
 lr 2.500000e-07 reg 1.000000e+06 batch 750 iters 2000 train accuracy: 0.404061 val accuracy: 0.386000
 lr 2.500000e-07 reg 2.500000e+06 batch 250 iters 250 train accuracy: 0.311673 val accuracy: 0.306000
 lr 2.500000e-07 reg 2.500000e+06 batch 250 iters 500 train accuracy: 0.286408 val accuracy: 0.301000
 lr 2.500000e-07 reg 2.500000e+06 batch 250 iters 1000 train accuracy: 0.313306 val accuracy: 0.326000
 lr 2.500000e-07 reg 2.500000e+06 batch 250 iters 2000 train accuracy: 0.298306 val accuracy: 0.277000
 lr 2.500000e-07 reg 2.500000e+06 batch 500 iters 250 train accuracy: 0.339673 val accuracy: 0.322000
 lr 2.500000e-07 reg 2.500000e+06 batch 500 iters 500 train accuracy: 0.342857 val accuracy: 0.355000
 lr 2.500000e-07 reg 2.500000e+06 batch 500 iters 1000 train accuracy: 0.324857 val accuracy: 0.304000
 lr 2.500000e-07 reg 2.500000e+06 batch 500 iters 2000 train accuracy: 0.351082 val accuracy: 0.344000
 lr 2.500000e-07 reg 2.500000e+06 batch 750 iters 250 train accuracy: 0.366163 val accuracy: 0.351000
 lr 2.500000e-07 reg 2.500000e+06 batch 750 iters 500 train accuracy: 0.375449 val accuracy: 0.382000
 lr 2.500000e-07 reg 2.500000e+06 batch 750 iters 1000 train accuracy: 0.359245 val accuracy: 0.353000
 lr 2.500000e-07 reg 2.500000e+06 batch 750 iters 2000 train accuracy: 0.374163 val accuracy: 0.359000
 lr 2.500000e-07 reg 5.000000e+06 batch 250 iters 250 train accuracy: 0.105061 val accuracy: 0.118000
 lr 2.500000e-07 reg 5.000000e+06 batch 250 iters 500 train accuracy: 0.104531 val accuracy: 0.116000
 lr 2.500000e-07 reg 5.000000e+06 batch 250 iters 1000 train accuracy: 0.084673 val accuracy: 0.091000
 lr 2.500000e-07 reg 5.000000e+06 batch 250 iters 2000 train accuracy: 0.100265 val accuracy: 0.087000
 lr 2.500000e-07 reg 5.000000e+06 batch 500 iters 250 train accuracy: 0.099306 val accuracy: 0.091000
 lr 2.500000e-07 reg 5.000000e+06 batch 500 iters 500 train accuracy: 0.098878 val accuracy: 0.105000
 lr 2.500000e-07 reg 5.000000e+06 batch 500 iters 1000 train accuracy: 0.103673 val accuracy: 0.110000
 lr 2.500000e-07 reg 5.000000e+06 batch 500 iters 2000 train accuracy: 0.100265 val accuracy: 0.087000
 lr 2.500000e-07 reg 5.000000e+06 batch 750 iters 250 train accuracy: 0.091224 val accuracy: 0.092000
 lr 2.500000e-07 reg 5.000000e+06 batch 750 iters 500 train accuracy: 0.082571 val accuracy: 0.081000
 lr 2.500000e-07 reg 5.000000e+06 batch 750 iters 1000 train accuracy: 0.104429 val accuracy: 0.095000
 lr 2.500000e-07 reg 5.000000e+06 batch 750 iters 2000 train accuracy: 0.100265 val accuracy: 0.087000
 lr 5.000000e-07 reg 1.000000e+01 batch 250 iters 250 train accuracy: 0.094694 val accuracy: 0.107000
 lr 5.000000e-07 reg 1.000000e+01 batch 250 iters 500 train accuracy: 0.126306 val accuracy: 0.110000
 lr 5.000000e-07 reg 1.000000e+01 batch 250 iters 1000 train accuracy: 0.165490 val accuracy: 0.172000
 lr 5.000000e-07 reg 1.000000e+01 batch 250 iters 2000 train accuracy: 0.221531 val accuracy: 0.183000
 lr 5.000000e-07 reg 1.000000e+01 batch 500 iters 250 train accuracy: 0.113204 val accuracy: 0.127000
 lr 5.000000e-07 reg 1.000000e+01 batch 500 iters 500 train accuracy: 0.121408 val accuracy: 0.126000
 lr 5.000000e-07 reg 1.000000e+01 batch 500 iters 1000 train accuracy: 0.152633 val accuracy: 0.163000
 lr 5.000000e-07 reg 1.000000e+01 batch 500 iters 2000 train accuracy: 0.220837 val accuracy: 0.212000
 lr 5.000000e-07 reg 1.000000e+01 batch 750 iters 250 train accuracy: 0.116653 val accuracy: 0.118000
 lr 5.000000e-07 reg 1.000000e+01 batch 750 iters 500 train accuracy: 0.141286 val accuracy: 0.161000
 lr 5.000000e-07 reg 1.000000e+01 batch 750 iters 1000 train accuracy: 0.170286 val accuracy: 0.183000

lr 5.000000e-07 reg 1.000000e+01 batch 750 iters 2000 train accuracy: 0.198449 val accuracy: 0.201000
 lr 5.000000e-07 reg 2.500000e+01 batch 250 iters 250 train accuracy: 0.092286 val accuracy: 0.104000
 lr 5.000000e-07 reg 2.500000e+01 batch 250 iters 500 train accuracy: 0.125490 val accuracy: 0.127000
 lr 5.000000e-07 reg 2.500000e+01 batch 250 iters 1000 train accuracy: 0.136755 val accuracy: 0.134000
 lr 5.000000e-07 reg 2.500000e+01 batch 250 iters 2000 train accuracy: 0.203490 val accuracy: 0.197000
 lr 5.000000e-07 reg 2.500000e+01 batch 500 iters 250 train accuracy: 0.107755 val accuracy: 0.120000
 lr 5.000000e-07 reg 2.500000e+01 batch 500 iters 500 train accuracy: 0.115878 val accuracy: 0.123000
 lr 5.000000e-07 reg 2.500000e+01 batch 500 iters 1000 train accuracy: 0.174306 val accuracy: 0.186000
 lr 5.000000e-07 reg 2.500000e+01 batch 500 iters 2000 train accuracy: 0.209776 val accuracy: 0.198000
 lr 5.000000e-07 reg 2.500000e+01 batch 750 iters 250 train accuracy: 0.132898 val accuracy: 0.130000
 lr 5.000000e-07 reg 2.500000e+01 batch 750 iters 500 train accuracy: 0.155041 val accuracy: 0.160000
 lr 5.000000e-07 reg 2.500000e+01 batch 750 iters 1000 train accuracy: 0.151449 val accuracy: 0.139000
 lr 5.000000e-07 reg 2.500000e+01 batch 750 iters 2000 train accuracy: 0.207163 val accuracy: 0.234000
 lr 5.000000e-07 reg 5.000000e+01 batch 250 iters 250 train accuracy: 0.130184 val accuracy: 0.136000
 lr 5.000000e-07 reg 5.000000e+01 batch 250 iters 500 train accuracy: 0.138592 val accuracy: 0.146000
 lr 5.000000e-07 reg 5.000000e+01 batch 250 iters 1000 train accuracy: 0.163143 val accuracy: 0.167000
 lr 5.000000e-07 reg 5.000000e+01 batch 250 iters 2000 train accuracy: 0.228694 val accuracy: 0.237000
 lr 5.000000e-07 reg 5.000000e+01 batch 500 iters 250 train accuracy: 0.098776 val accuracy: 0.103000
 lr 5.000000e-07 reg 5.000000e+01 batch 500 iters 500 train accuracy: 0.116898 val accuracy: 0.115000
 lr 5.000000e-07 reg 5.000000e+01 batch 500 iters 1000 train accuracy: 0.143857 val accuracy: 0.142000
 lr 5.000000e-07 reg 5.000000e+01 batch 500 iters 2000 train accuracy: 0.204939 val accuracy: 0.193000
 lr 5.000000e-07 reg 5.000000e+01 batch 750 iters 250 train accuracy: 0.097551 val accuracy: 0.098000
 lr 5.000000e-07 reg 5.000000e+01 batch 750 iters 500 train accuracy: 0.146490 val accuracy: 0.143000
 lr 5.000000e-07 reg 5.000000e+01 batch 750 iters 1000 train accuracy: 0.141898 val accuracy: 0.133000
 lr 5.000000e-07 reg 5.000000e+01 batch 750 iters 2000 train accuracy: 0.219592 val accuracy: 0.220000
 lr 5.000000e-07 reg 1.000000e+02 batch 250 iters 250 train accuracy: 0.125694 val accuracy: 0.141000
 lr 5.000000e-07 reg 1.000000e+02 batch 250 iters 500 train accuracy: 0.140265 val accuracy: 0.134000
 lr 5.000000e-07 reg 1.000000e+02 batch 250 iters 1000 train accuracy: 0.162939 val accuracy: 0.186000
 lr 5.000000e-07 reg 1.000000e+02 batch 250 iters 2000 train accuracy: 0.223082 val accuracy: 0.221000
 lr 5.000000e-07 reg 1.000000e+02 batch 500 iters 250 train accuracy: 0.132000 val accuracy: 0.146000
 lr 5.000000e-07 reg 1.000000e+02 batch 500 iters 500 train accuracy: 0.143265 val accuracy: 0.179000
 lr 5.000000e-07 reg 1.000000e+02 batch 500 iters 1000 train accuracy: 0.166184 val accuracy: 0.155000
 lr 5.000000e-07 reg 1.000000e+02 batch 500 iters 2000 train accuracy: 0.217408 val accuracy: 0.222000
 lr 5.000000e-07 reg 1.000000e+02 batch 750 iters 250 train accuracy: 0.107020 val accuracy: 0.113000
 lr 5.000000e-07 reg 1.000000e+02 batch 750 iters 500 train accuracy: 0.124592 val accuracy: 0.108000
 lr 5.000000e-07 reg 1.000000e+02 batch 750 iters 1000 train accuracy: 0.149898 val accuracy: 0.141000
 lr 5.000000e-07 reg 1.000000e+02 batch 750 iters 2000 train accuracy: 0.217245 val accuracy: 0.216000
 lr 5.000000e-07 reg 2.500000e+02 batch 250 iters 250 train accuracy: 0.088735 val accuracy: 0.100000
 lr 5.000000e-07 reg 2.500000e+02 batch 250 iters 500 train accuracy: 0.133633 val accuracy: 0.130000
 lr 5.000000e-07 reg 2.500000e+02 batch 250 iters 1000 train accuracy: 0.165653 val accuracy: 0.187000
 lr 5.000000e-07 reg 2.500000e+02 batch 250 iters 2000 train accuracy: 0.237367 val accuracy: 0.234000
 lr 5.000000e-07 reg 2.500000e+02 batch 500 iters 250 train accuracy: 0.103122 val accuracy: 0.095000
 lr 5.000000e-07 reg 2.500000e+02 batch 500 iters 500 train accuracy: 0.141653 val accuracy: 0.153000
 lr 5.000000e-07 reg 2.500000e+02 batch 500 iters 1000 train accuracy: 0.163898 val accuracy: 0.151000
 lr 5.000000e-07 reg 2.500000e+02 batch 500 iters 2000 train accuracy: 0.254286 val accuracy: 0.248000
 lr 5.000000e-07 reg 2.500000e+02 batch 750 iters 250 train accuracy: 0.107122 val accuracy: 0.113000
 lr 5.000000e-07 reg 2.500000e+02 batch 750 iters 500 train accuracy: 0.097449 val accuracy: 0.107000
 lr 5.000000e-07 reg 2.500000e+02 batch 750 iters 1000 train accuracy: 0.141224 val accuracy: 0.143000
 lr 5.000000e-07 reg 2.500000e+02 batch 750 iters 2000 train accuracy: 0.238837 val accuracy: 0.221000
 lr 5.000000e-07 reg 5.000000e+02 batch 250 iters 250 train accuracy: 0.121000 val accuracy: 0.124000
 lr 5.000000e-07 reg 5.000000e+02 batch 250 iters 500 train accuracy: 0.135776 val accuracy: 0.128000
 lr 5.000000e-07 reg 5.000000e+02 batch 250 iters 1000 train accuracy: 0.170020 val accuracy: 0.189000
 lr 5.000000e-07 reg 5.000000e+02 batch 250 iters 2000 train accuracy: 0.269510 val accuracy: 0.274000
 lr 5.000000e-07 reg 5.000000e+02 batch 500 iters 250 train accuracy: 0.080653 val accuracy: 0.075000
 lr 5.000000e-07 reg 5.000000e+02 batch 500 iters 500 train accuracy: 0.127061 val accuracy: 0.141000
 lr 5.000000e-07 reg 5.000000e+02 batch 500 iters 1000 train accuracy: 0.181224 val accuracy: 0.157000
 lr 5.000000e-07 reg 5.000000e+02 batch 500 iters 2000 train accuracy: 0.260327 val accuracy: 0.254000
 lr 5.000000e-07 reg 5.000000e+02 batch 750 iters 250 train accuracy: 0.120816 val accuracy: 0.121000
 lr 5.000000e-07 reg 5.000000e+02 batch 750 iters 500 train accuracy: 0.143673 val accuracy: 0.128000
 lr 5.000000e-07 reg 5.000000e+02 batch 750 iters 1000 train accuracy: 0.171837 val accuracy: 0.195000
 lr 5.000000e-07 reg 5.000000e+02 batch 750 iters 2000 train accuracy: 0.254347 val accuracy: 0.284000
 lr 5.000000e-07 reg 1.000000e+03 batch 250 iters 250 train accuracy: 0.095429 val accuracy: 0.091000
 lr 5.000000e-07 reg 1.000000e+03 batch 250 iters 500 train accuracy: 0.123918 val accuracy: 0.125000
 lr 5.000000e-07 reg 1.000000e+03 batch 250 iters 1000 train accuracy: 0.192755 val accuracy: 0.185000
 lr 5.000000e-07 reg 1.000000e+03 batch 250 iters 2000 train accuracy: 0.334816 val accuracy: 0.354000
 lr 5.000000e-07 reg 1.000000e+03 batch 500 iters 250 train accuracy: 0.101755 val accuracy: 0.100000
 lr 5.000000e-07 reg 1.000000e+03 batch 500 iters 500 train accuracy: 0.153980 val accuracy: 0.150000
 lr 5.000000e-07 reg 1.000000e+03 batch 500 iters 1000 train accuracy: 0.199469 val accuracy: 0.220000

1r 5.000000e-07	reg	1.000000e+03	batch	500	iters	2000	train accuracy: 0.349571	val accuracy: 0.357000
1r 5.000000e-07	reg	1.000000e+03	batch	750	iters	250	train accuracy: 0.106265	val accuracy: 0.114000
1r 5.000000e-07	reg	1.000000e+03	batch	750	iters	500	train accuracy: 0.105429	val accuracy: 0.100000
1r 5.000000e-07	reg	1.000000e+03	batch	750	iters	1000	train accuracy: 0.204122	val accuracy: 0.186000
1r 5.000000e-07	reg	1.000000e+03	batch	750	iters	2000	train accuracy: 0.340082	val accuracy: 0.339000
1r 5.000000e-07	reg	2.500000e+03	batch	250	iters	250	train accuracy: 0.134020	val accuracy: 0.136000
1r 5.000000e-07	reg	2.500000e+03	batch	250	iters	500	train accuracy: 0.142347	val accuracy: 0.146000
1r 5.000000e-07	reg	2.500000e+03	batch	250	iters	1000	train accuracy: 0.289061	val accuracy: 0.280000
1r 5.000000e-07	reg	2.500000e+03	batch	250	iters	2000	train accuracy: 0.414980	val accuracy: 0.425000
1r 5.000000e-07	reg	2.500000e+03	batch	500	iters	250	train accuracy: 0.108714	val accuracy: 0.089000
1r 5.000000e-07	reg	2.500000e+03	batch	500	iters	500	train accuracy: 0.147571	val accuracy: 0.152000
1r 5.000000e-07	reg	2.500000e+03	batch	500	iters	1000	train accuracy: 0.287898	val accuracy: 0.285000
1r 5.000000e-07	reg	2.500000e+03	batch	500	iters	2000	train accuracy: 0.417469	val accuracy: 0.420000
1r 5.000000e-07	reg	2.500000e+03	batch	750	iters	250	train accuracy: 0.129755	val accuracy: 0.120000
1r 5.000000e-07	reg	2.500000e+03	batch	750	iters	500	train accuracy: 0.173327	val accuracy: 0.202000
1r 5.000000e-07	reg	2.500000e+03	batch	750	iters	1000	train accuracy: 0.296980	val accuracy: 0.316000
1r 5.000000e-07	reg	2.500000e+03	batch	750	iters	2000	train accuracy: 0.413755	val accuracy: 0.412000
1r 5.000000e-07	reg	5.000000e+03	batch	250	iters	250	train accuracy: 0.121449	val accuracy: 0.105000
1r 5.000000e-07	reg	5.000000e+03	batch	250	iters	500	train accuracy: 0.244776	val accuracy: 0.250000
1r 5.000000e-07	reg	5.000000e+03	batch	250	iters	1000	train accuracy: 0.411408	val accuracy: 0.418000
1r 5.000000e-07	reg	5.000000e+03	batch	250	iters	2000	train accuracy: 0.416163	val accuracy: 0.414000
1r 5.000000e-07	reg	5.000000e+03	batch	500	iters	250	train accuracy: 0.143490	val accuracy: 0.156000
1r 5.000000e-07	reg	5.000000e+03	batch	500	iters	500	train accuracy: 0.212000	val accuracy: 0.185000
1r 5.000000e-07	reg	5.000000e+03	batch	500	iters	1000	train accuracy: 0.407408	val accuracy: 0.398000
1r 5.000000e-07	reg	5.000000e+03	batch	500	iters	2000	train accuracy: 0.415429	val accuracy: 0.420000
1r 5.000000e-07	reg	5.000000e+03	batch	750	iters	250	train accuracy: 0.098673	val accuracy: 0.099000
1r 5.000000e-07	reg	5.000000e+03	batch	750	iters	500	train accuracy: 0.224286	val accuracy: 0.228000
1r 5.000000e-07	reg	5.000000e+03	batch	750	iters	1000	train accuracy: 0.416143	val accuracy: 0.418000
1r 5.000000e-07	reg	5.000000e+03	batch	750	iters	2000	train accuracy: 0.416306	val accuracy: 0.421000
1r 5.000000e-07	reg	1.000000e+04	batch	250	iters	250	train accuracy: 0.161347	val accuracy: 0.152000
1r 5.000000e-07	reg	1.000000e+04	batch	250	iters	500	train accuracy: 0.400082	val accuracy: 0.414000
1r 5.000000e-07	reg	1.000000e+04	batch	250	iters	1000	train accuracy: 0.417755	val accuracy: 0.420000
1r 5.000000e-07	reg	1.000000e+04	batch	250	iters	2000	train accuracy: 0.414653	val accuracy: 0.410000
1r 5.000000e-07	reg	1.000000e+04	batch	500	iters	250	train accuracy: 0.170286	val accuracy: 0.199000
1r 5.000000e-07	reg	1.000000e+04	batch	500	iters	500	train accuracy: 0.398592	val accuracy: 0.398000
1r 5.000000e-07	reg	1.000000e+04	batch	500	iters	1000	train accuracy: 0.415939	val accuracy: 0.421000
1r 5.000000e-07	reg	1.000000e+04	batch	500	iters	2000	train accuracy: 0.413408	val accuracy: 0.421000
1								

1r 5.000000e-07	reg	1.000000e+05	batch	250	iters	2000	train accuracy: 0.408735	val accuracy: 0.405000
1r 5.000000e-07	reg	1.000000e+05	batch	500	iters	250	train accuracy: 0.405571	val accuracy: 0.390000
1r 5.000000e-07	reg	1.000000e+05	batch	500	iters	500	train accuracy: 0.406816	val accuracy: 0.404000
1r 5.000000e-07	reg	1.000000e+05	batch	500	iters	1000	train accuracy: 0.408776	val accuracy: 0.416000
1r 5.000000e-07	reg	1.000000e+05	batch	500	iters	2000	train accuracy: 0.416265	val accuracy: 0.412000
1r 5.000000e-07	reg	1.000000e+05	batch	750	iters	250	train accuracy: 0.415551	val accuracy: 0.413000
1r 5.000000e-07	reg	1.000000e+05	batch	750	iters	500	train accuracy: 0.415653	val accuracy: 0.408000
1r 5.000000e-07	reg	1.000000e+05	batch	750	iters	1000	train accuracy: 0.412531	val accuracy: 0.420000
1r 5.000000e-07	reg	1.000000e+05	batch	750	iters	2000	train accuracy: 0.414286	val accuracy: 0.425000
1r 5.000000e-07	reg	2.500000e+05	batch	250	iters	250	train accuracy: 0.394714	val accuracy: 0.398000
1r 5.000000e-07	reg	2.500000e+05	batch	250	iters	500	train accuracy: 0.407816	val accuracy: 0.408000
1r 5.000000e-07	reg	2.500000e+05	batch	250	iters	1000	train accuracy: 0.387714	val accuracy: 0.403000
1r 5.000000e-07	reg	2.500000e+05	batch	250	iters	2000	train accuracy: 0.415143	val accuracy: 0.431000
1r 5.000000e-07	reg	2.500000e+05	batch	500	iters	250	train accuracy: 0.409918	val accuracy: 0.425000
1r 5.000000e-07	reg	2.500000e+05	batch	500	iters	500	train accuracy: 0.395490	val accuracy: 0.385000
1r 5.000000e-07	reg	2.500000e+05	batch	500	iters	1000	train accuracy: 0.403020	val accuracy: 0.400000
1r 5.000000e-07	reg	2.500000e+05	batch	500	iters	2000	train accuracy: 0.401816	val accuracy: 0.415000
1r 5.000000e-07	reg	2.500000e+05	batch	750	iters	250	train accuracy: 0.401347	val accuracy: 0.387000
1r 5.000000e-07	reg	2.500000e+05	batch	750	iters	500	train accuracy: 0.398571	val accuracy: 0.387000
1r 5.000000e-07	reg	2.500000e+05	batch	750	iters	1000	train accuracy: 0.405388	val accuracy: 0.406000
1r 5.000000e-07	reg	2.500000e+05	batch	750	iters	2000	train accuracy: 0.399224	val accuracy: 0.404000
1r 5.000000e-07	reg	5.000000e+05	batch	250	iters	250	train accuracy: 0.371061	val accuracy: 0.363000
1r 5.000000e-07	reg	5.000000e+05	batch	250	iters	500	train accuracy: 0.372939	val accuracy: 0.374000
1r 5.000000e-07	reg	5.000000e+05	batch	250	iters	1000	train accuracy: 0.386408	val accuracy: 0.367000
1r 5.000000e-07	reg	5.000000e+05	batch	250	iters	2000	train accuracy: 0.387959	val accuracy: 0.388000
1r 5.000000e-07	reg	5.000000e+05	batch	500	iters	250	train accuracy: 0.386490	val accuracy: 0.371000
1r 5.000000e-07	reg	5.000000e+05	batch	500	iters	500	train accuracy: 0.391653	val accuracy: 0.403000
1r 5.000000e-07	reg	5.000000e+05	batch	500	iters	1000	train accuracy: 0.407429	val accuracy: 0.399000
1r 5.000000e-07	reg	5.000000e+05	batch	500	iters	2000	train accuracy: 0.385143	val accuracy: 0.367000
1r 5.000000e-07	reg	5.000000e+05	batch	750	iters	250	train accuracy: 0.397061	val accuracy: 0.398000
1r 5.000000e-07	reg	5.000000e+05	batch	750	iters	500	train accuracy: 0.404551	val accuracy: 0.413000
1r 5.000000e-07	reg	5.000000e+05	batch	750	iters	1000	train accuracy: 0.410204	val accuracy: 0.414000
1r 5.000000e-07	reg	5.000000e+05	batch	750	iters	2000	train accuracy: 0.400837	val accuracy: 0.407000
1r 5.000000e-07	reg	1.000000e+06	batch	250	iters	250	train accuracy: 0.325551	val accuracy: 0.315000
1r 5.000000e-07	reg	1.000000e+06	batch	250	iters	500	train accuracy: 0.333265	val accuracy: 0.297000
1r 5.000000e-07	reg	1.000000e+06	batch	250	iters	1000	train accuracy: 0.340082	val accuracy: 0.352000
1r 5.000000e-07	reg	1.000000e+06	batch	250	iters	2000	train accuracy: 0.333531	val accuracy: 0.327000
1								

lr 5.000000e-07 reg 5.000000e+06 batch 750 iters 2000 train accuracy: 0.100265 val accuracy: 0.087000
 lr 1.000000e-06 reg 1.000000e+01 batch 250 iters 250 train accuracy: 0.117327 val accuracy: 0.105000
 lr 1.000000e-06 reg 1.000000e+01 batch 250 iters 500 train accuracy: 0.157898 val accuracy: 0.173000
 lr 1.000000e-06 reg 1.000000e+01 batch 250 iters 1000 train accuracy: 0.220673 val accuracy: 0.218000
 lr 1.000000e-06 reg 1.000000e+01 batch 250 iters 2000 train accuracy: 0.292265 val accuracy: 0.305000
 lr 1.000000e-06 reg 1.000000e+01 batch 500 iters 250 train accuracy: 0.121612 val accuracy: 0.125000
 lr 1.000000e-06 reg 1.000000e+01 batch 500 iters 500 train accuracy: 0.125653 val accuracy: 0.116000
 lr 1.000000e-06 reg 1.000000e+01 batch 500 iters 1000 train accuracy: 0.233020 val accuracy: 0.226000
 lr 1.000000e-06 reg 1.000000e+01 batch 500 iters 2000 train accuracy: 0.302224 val accuracy: 0.301000
 lr 1.000000e-06 reg 1.000000e+01 batch 750 iters 250 train accuracy: 0.145367 val accuracy: 0.150000
 lr 1.000000e-06 reg 1.000000e+01 batch 750 iters 500 train accuracy: 0.175531 val accuracy: 0.186000
 lr 1.000000e-06 reg 1.000000e+01 batch 750 iters 1000 train accuracy: 0.213347 val accuracy: 0.215000
 lr 1.000000e-06 reg 1.000000e+01 batch 750 iters 2000 train accuracy: 0.282184 val accuracy: 0.256000
 lr 1.000000e-06 reg 2.500000e+01 batch 250 iters 250 train accuracy: 0.099857 val accuracy: 0.118000
 lr 1.000000e-06 reg 2.500000e+01 batch 250 iters 500 train accuracy: 0.163816 val accuracy: 0.165000
 lr 1.000000e-06 reg 2.500000e+01 batch 250 iters 1000 train accuracy: 0.233122 val accuracy: 0.239000
 lr 1.000000e-06 reg 2.500000e+01 batch 250 iters 2000 train accuracy: 0.300612 val accuracy: 0.311000
 lr 1.000000e-06 reg 2.500000e+01 batch 500 iters 250 train accuracy: 0.117735 val accuracy: 0.112000
 lr 1.000000e-06 reg 2.500000e+01 batch 500 iters 500 train accuracy: 0.141551 val accuracy: 0.160000
 lr 1.000000e-06 reg 2.500000e+01 batch 500 iters 1000 train accuracy: 0.223204 val accuracy: 0.210000
 lr 1.000000e-06 reg 2.500000e+01 batch 500 iters 2000 train accuracy: 0.295510 val accuracy: 0.293000
 lr 1.000000e-06 reg 2.500000e+01 batch 750 iters 250 train accuracy: 0.128286 val accuracy: 0.120000
 lr 1.000000e-06 reg 2.500000e+01 batch 750 iters 500 train accuracy: 0.178959 val accuracy: 0.177000
 lr 1.000000e-06 reg 2.500000e+01 batch 750 iters 1000 train accuracy: 0.213449 val accuracy: 0.214000
 lr 1.000000e-06 reg 2.500000e+01 batch 750 iters 2000 train accuracy: 0.283735 val accuracy: 0.294000
 lr 1.000000e-06 reg 5.000000e+01 batch 250 iters 250 train accuracy: 0.123102 val accuracy: 0.128000
 lr 1.000000e-06 reg 5.000000e+01 batch 250 iters 500 train accuracy: 0.155163 val accuracy: 0.138000
 lr 1.000000e-06 reg 5.000000e+01 batch 250 iters 1000 train accuracy: 0.228878 val accuracy: 0.225000
 lr 1.000000e-06 reg 5.000000e+01 batch 250 iters 2000 train accuracy: 0.282939 val accuracy: 0.291000
 lr 1.000000e-06 reg 5.000000e+01 batch 500 iters 250 train accuracy: 0.118082 val accuracy: 0.104000
 lr 1.000000e-06 reg 5.000000e+01 batch 500 iters 500 train accuracy: 0.166265 val accuracy: 0.170000
 lr 1.000000e-06 reg 5.000000e+01 batch 500 iters 1000 train accuracy: 0.224143 val accuracy: 0.227000
 lr 1.000000e-06 reg 5.000000e+01 batch 500 iters 2000 train accuracy: 0.304592 val accuracy: 0.299000
 lr 1.000000e-06 reg 5.000000e+01 batch 750 iters 250 train accuracy: 0.117980 val accuracy: 0.110000
 lr 1.000000e-06 reg 5.000000e+01 batch 750 iters 500 train accuracy: 0.146102 val accuracy: 0.157000
 lr 1.000000e-06 reg 5.000000e+01 batch 750 iters 1000 train accuracy: 0.219163 val accuracy: 0.231000
 lr 1.000000e-06 reg 5.000000e+01 batch 750 iters 2000 train accuracy: 0.286327 val accuracy: 0.280000
 lr 1.000000e-06 reg 1.000000e+02 batch 250 iters 250 train accuracy: 0.140449 val accuracy: 0.145000
 lr 1.000000e-06 reg 1.000000e+02 batch 250 iters 500 train accuracy: 0.137776 val accuracy: 0.138000
 lr 1.000000e-06 reg 1.000000e+02 batch 250 iters 1000 train accuracy: 0.238755 val accuracy: 0.226000
 lr 1.000000e-06 reg 1.000000e+02 batch 250 iters 2000 train accuracy: 0.309163 val accuracy: 0.315000
 lr 1.000000e-06 reg 1.000000e+02 batch 500 iters 250 train accuracy: 0.143061 val accuracy: 0.158000
 lr 1.000000e-06 reg 1.000000e+02 batch 500 iters 500 train accuracy: 0.165286 val accuracy: 0.171000
 lr 1.000000e-06 reg 1.000000e+02 batch 500 iters 1000 train accuracy: 0.208571 val accuracy: 0.212000
 lr 1.000000e-06 reg 1.000000e+02 batch 500 iters 2000 train accuracy: 0.308531 val accuracy: 0.329000
 lr 1.000000e-06 reg 1.000000e+02 batch 750 iters 250 train accuracy: 0.150571 val accuracy: 0.145000
 lr 1.000000e-06 reg 1.000000e+02 batch 750 iters 500 train accuracy: 0.168837 val accuracy: 0.168000
 lr 1.000000e-06 reg 1.000000e+02 batch 750 iters 1000 train accuracy: 0.208041 val accuracy: 0.229000
 lr 1.000000e-06 reg 1.000000e+02 batch 750 iters 2000 train accuracy: 0.301061 val accuracy: 0.308000
 lr 1.000000e-06 reg 2.500000e+02 batch 250 iters 250 train accuracy: 0.125673 val accuracy: 0.122000
 lr 1.000000e-06 reg 2.500000e+02 batch 250 iters 500 train accuracy: 0.180224 val accuracy: 0.178000
 lr 1.000000e-06 reg 2.500000e+02 batch 250 iters 1000 train accuracy: 0.252878 val accuracy: 0.242000
 lr 1.000000e-06 reg 2.500000e+02 batch 250 iters 2000 train accuracy: 0.338816 val accuracy: 0.306000
 lr 1.000000e-06 reg 2.500000e+02 batch 500 iters 250 train accuracy: 0.123408 val accuracy: 0.132000
 lr 1.000000e-06 reg 2.500000e+02 batch 500 iters 500 train accuracy: 0.166592 val accuracy: 0.162000
 lr 1.000000e-06 reg 2.500000e+02 batch 500 iters 1000 train accuracy: 0.241245 val accuracy: 0.247000
 lr 1.000000e-06 reg 2.500000e+02 batch 500 iters 2000 train accuracy: 0.337000 val accuracy: 0.345000
 lr 1.000000e-06 reg 2.500000e+02 batch 750 iters 250 train accuracy: 0.118388 val accuracy: 0.122000
 lr 1.000000e-06 reg 2.500000e+02 batch 750 iters 500 train accuracy: 0.166449 val accuracy: 0.176000
 lr 1.000000e-06 reg 2.500000e+02 batch 750 iters 1000 train accuracy: 0.228612 val accuracy: 0.226000
 lr 1.000000e-06 reg 2.500000e+02 batch 750 iters 2000 train accuracy: 0.349449 val accuracy: 0.342000
 lr 1.000000e-06 reg 5.000000e+02 batch 250 iters 250 train accuracy: 0.175204 val accuracy: 0.167000
 lr 1.000000e-06 reg 5.000000e+02 batch 250 iters 500 train accuracy: 0.163653 val accuracy: 0.160000
 lr 1.000000e-06 reg 5.000000e+02 batch 250 iters 1000 train accuracy: 0.257939 val accuracy: 0.262000
 lr 1.000000e-06 reg 5.000000e+02 batch 250 iters 2000 train accuracy: 0.386204 val accuracy: 0.374000
 lr 1.000000e-06 reg 5.000000e+02 batch 500 iters 250 train accuracy: 0.128224 val accuracy: 0.128000
 lr 1.000000e-06 reg 5.000000e+02 batch 500 iters 500 train accuracy: 0.171224 val accuracy: 0.181000
 lr 1.000000e-06 reg 5.000000e+02 batch 500 iters 1000 train accuracy: 0.267959 val accuracy: 0.260000

1r	1.000000e-06	reg	5.000000e+02	batch	500	iters	2000	train accuracy: 0.386306	val accuracy: 0.402000
1r	1.000000e-06	reg	5.000000e+02	batch	750	iters	250	train accuracy: 0.142918	val accuracy: 0.128000
1r	1.000000e-06	reg	5.000000e+02	batch	750	iters	500	train accuracy: 0.200796	val accuracy: 0.177000
1r	1.000000e-06	reg	5.000000e+02	batch	750	iters	1000	train accuracy: 0.292531	val accuracy: 0.304000
1r	1.000000e-06	reg	5.000000e+02	batch	750	iters	2000	train accuracy: 0.383735	val accuracy: 0.406000
1r	1.000000e-06	reg	1.000000e+03	batch	250	iters	250	train accuracy: 0.147204	val accuracy: 0.123000
1r	1.000000e-06	reg	1.000000e+03	batch	250	iters	500	train accuracy: 0.202082	val accuracy: 0.204000
1r	1.000000e-06	reg	1.000000e+03	batch	250	iters	1000	train accuracy: 0.336000	val accuracy: 0.360000
1r	1.000000e-06	reg	1.000000e+03	batch	250	iters	2000	train accuracy: 0.411959	val accuracy: 0.429000
1r	1.000000e-06	reg	1.000000e+03	batch	500	iters	250	train accuracy: 0.142347	val accuracy: 0.128000
1r	1.000000e-06	reg	1.000000e+03	batch	500	iters	500	train accuracy: 0.191939	val accuracy: 0.195000
1r	1.000000e-06	reg	1.000000e+03	batch	500	iters	1000	train accuracy: 0.323347	val accuracy: 0.317000
1r	1.000000e-06	reg	1.000000e+03	batch	500	iters	2000	train accuracy: 0.414163	val accuracy: 0.421000
1r	1.000000e-06	reg	1.000000e+03	batch	750	iters	250	train accuracy: 0.129429	val accuracy: 0.122000
1r	1.000000e-06	reg	1.000000e+03	batch	750	iters	500	train accuracy: 0.188490	val accuracy: 0.212000
1r	1.000000e-06	reg	1.000000e+03	batch	750	iters	1000	train accuracy: 0.334653	val accuracy: 0.316000
1r	1.000000e-06	reg	1.000000e+03	batch	750	iters	2000	train accuracy: 0.417959	val accuracy: 0.420000
1r	1.000000e-06	reg	2.500000e+03	batch	250	iters	250	train accuracy: 0.146429	val accuracy: 0.139000
1r	1.000000e-06	reg	2.500000e+03	batch	250	iters	500	train accuracy: 0.296898	val accuracy: 0.278000
1r	1.000000e-06	reg	2.500000e+03	batch	250	iters	1000	train accuracy: 0.414918	val accuracy: 0.428000
1r	1.000000e-06	reg	2.500000e+03	batch	250	iters	2000	train accuracy: 0.414837	val accuracy: 0.418000
1r	1.000000e-06	reg	2.500000e+03	batch	500	iters	250	train accuracy: 0.145776	val accuracy: 0.148000
1r	1.000000e-06	reg	2.500000e+03	batch	500	iters	500	train accuracy: 0.308694	val accuracy: 0.320000
1r	1.000000e-06	reg	2.500000e+03	batch	500	iters	1000	train accuracy: 0.412939	val accuracy: 0.410000
1r	1.000000e-06	reg	2.500000e+03	batch	500	iters	2000	train accuracy: 0.414735	val accuracy: 0.420000
1r	1.000000e-06	reg	2.500000e+03	batch	750	iters	250	train accuracy: 0.159020	val accuracy: 0.157000
1r	1.000000e-06	reg	2.500000e+03	batch	750	iters	500	train accuracy: 0.297469	val accuracy: 0.294000
1r	1.000000e-06	reg	2.500000e+03	batch	750	iters	1000	train accuracy: 0.412510	val accuracy: 0.407000
1r	1.000000e-06	reg	2.500000e+03	batch	750	iters	2000	train accuracy: 0.415653	val accuracy: 0.416000
1r	1.000000e-06	reg	5.000000e+03	batch	250	iters	250	train accuracy: 0.216367	val accuracy: 0.208000
1r	1.000000e-06	reg	5.000000e+03	batch	250	iters	500	train accuracy: 0.412939	val accuracy: 0.422000
1r	1.000000e-06	reg	5.000000e+03	batch	250	iters	1000	train accuracy: 0.410980	val accuracy: 0.413000
1r	1.000000e-06	reg	5.000000e+03	batch	250	iters	2000	train accuracy: 0.416041	val accuracy: 0.422000
1r	1.000000e-06	reg	5.000000e+03	batch	500	iters	250	train accuracy: 0.204245	val accuracy: 0.191000
1r	1.000000e-06	reg	5.000000e+03	batch	500	iters	500	train accuracy: 0.404816	val accuracy: 0.411000
1r	1.000000e-06	reg	5.000000e+03	batch	500	iters	1000	train accuracy: 0.	

lr	1.000000e-06	reg	5.000000e+04	batch	250	iters	2000	train accuracy: 0.408959	val accuracy: 0.409000
lr	1.000000e-06	reg	5.000000e+04	batch	500	iters	250	train accuracy: 0.409939	val accuracy: 0.413000
lr	1.000000e-06	reg	5.000000e+04	batch	500	iters	500	train accuracy: 0.407306	val accuracy: 0.399000
lr	1.000000e-06	reg	5.000000e+04	batch	500	iters	1000	train accuracy: 0.410429	val accuracy: 0.417000
lr	1.000000e-06	reg	5.000000e+04	batch	500	iters	2000	train accuracy: 0.413816	val accuracy: 0.394000
lr	1.000000e-06	reg	5.000000e+04	batch	750	iters	250	train accuracy: 0.406265	val accuracy: 0.404000
lr	1.000000e-06	reg	5.000000e+04	batch	750	iters	500	train accuracy: 0.415653	val accuracy: 0.425000
lr	1.000000e-06	reg	5.000000e+04	batch	750	iters	1000	train accuracy: 0.409224	val accuracy: 0.400000
lr	1.000000e-06	reg	5.000000e+04	batch	750	iters	2000	train accuracy: 0.414102	val accuracy: 0.422000
lr	1.000000e-06	reg	1.000000e+05	batch	250	iters	250	train accuracy: 0.398735	val accuracy: 0.403000
lr	1.000000e-06	reg	1.000000e+05	batch	250	iters	500	train accuracy: 0.407755	val accuracy: 0.402000
lr	1.000000e-06	reg	1.000000e+05	batch	250	iters	1000	train accuracy: 0.397184	val accuracy: 0.405000
lr	1.000000e-06	reg	1.000000e+05	batch	250	iters	2000	train accuracy: 0.404673	val accuracy: 0.406000
lr	1.000000e-06	reg	1.000000e+05	batch	500	iters	250	train accuracy: 0.407020	val accuracy: 0.420000
lr	1.000000e-06	reg	1.000000e+05	batch	500	iters	500	train accuracy: 0.401592	val accuracy: 0.406000
lr	1.000000e-06	reg	1.000000e+05	batch	500	iters	1000	train accuracy: 0.402449	val accuracy: 0.409000
lr	1.000000e-06	reg	1.000000e+05	batch	500	iters	2000	train accuracy: 0.403388	val accuracy: 0.399000
lr	1.000000e-06	reg	1.000000e+05	batch	750	iters	250	train accuracy: 0.411306	val accuracy: 0.419000
lr	1.000000e-06	reg	1.000000e+05	batch	750	iters	500	train accuracy: 0.414041	val accuracy: 0.425000
lr	1.000000e-06	reg	1.000000e+05	batch	750	iters	1000	train accuracy: 0.410245	val accuracy: 0.395000
lr	1.000000e-06	reg	1.000000e+05	batch	750	iters	2000	train accuracy: 0.417449	val accuracy: 0.425000
lr	1.000000e-06	reg	2.500000e+05	batch	250	iters	250	train accuracy: 0.377735	val accuracy: 0.373000
lr	1.000000e-06	reg	2.500000e+05	batch	250	iters	500	train accuracy: 0.368265	val accuracy: 0.368000
lr	1.000000e-06	reg	2.500000e+05	batch	250	iters	1000	train accuracy: 0.381224	val accuracy: 0.387000
lr	1.000000e-06	reg	2.500000e+05	batch	250	iters	2000	train accuracy: 0.380653	val accuracy: 0.366000
lr	1.000000e-06	reg	2.500000e+05	batch	500	iters	250	train accuracy: 0.390694	val accuracy: 0.390000
lr	1.000000e-06	reg	2.500000e+05	batch	500	iters	500	train accuracy: 0.395776	val accuracy: 0.397000
lr	1.000000e-06	reg	2.500000e+05	batch	500	iters	1000	train accuracy: 0.372755	val accuracy: 0.379000
lr	1.000000e-06	reg	2.500000e+05	batch	500	iters	2000	train accuracy: 0.392000	val accuracy: 0.406000
lr	1.000000e-06	reg	2.500000e+05	batch	750	iters	250	train accuracy: 0.397163	val accuracy: 0.386000
lr	1.000000e-06	reg	2.500000e+05	batch	750	iters	500	train accuracy: 0.395224	val accuracy: 0.384000
lr	1.000000e-06	reg	2.500000e+05	batch	750	iters	1000	train accuracy: 0.402837	val accuracy: 0.381000
lr	1.000000e-06	reg	2.500000e+05	batch	750	iters	2000	train accuracy: 0.398673	val accuracy: 0.397000
lr	1.000000e-06	reg	5.000000e+05	batch	250	iters	250	train accuracy: 0.317061	val accuracy: 0.315000
lr	1.000000e-06	reg	5.000000e+05	batch	250	iters	500	train accuracy: 0.331939	val accuracy: 0.317000
lr	1.000000e-06	reg	5.000000e+05	batch	250	iters	1000	train accuracy: 0.338878	val accuracy: 0.375000
lr									

lr 1.000000e-06	reg	2.500000e+06	batch	750	iters	2000	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-06	reg	5.000000e+06	batch	250	iters	250	train accuracy: 0.107469	val accuracy: 0.101000
lr 1.000000e-06	reg	5.000000e+06	batch	250	iters	500	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-06	reg	5.000000e+06	batch	250	iters	1000	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-06	reg	5.000000e+06	batch	250	iters	2000	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-06	reg	5.000000e+06	batch	500	iters	250	train accuracy: 0.111102	val accuracy: 0.109000
lr 1.000000e-06	reg	5.000000e+06	batch	500	iters	500	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-06	reg	5.000000e+06	batch	500	iters	1000	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-06	reg	5.000000e+06	batch	500	iters	2000	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-06	reg	5.000000e+06	batch	750	iters	250	train accuracy: 0.096082	val accuracy: 0.097000
lr 1.000000e-06	reg	5.000000e+06	batch	750	iters	500	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-06	reg	5.000000e+06	batch	750	iters	1000	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-06	reg	5.000000e+06	batch	750	iters	2000	train accuracy: 0.100265	val accuracy: 0.087000
lr 2.500000e-06	reg	1.000000e+01	batch	250	iters	250	train accuracy: 0.197592	val accuracy: 0.179000
lr 2.500000e-06	reg	1.000000e+01	batch	250	iters	500	train accuracy: 0.240939	val accuracy: 0.268000
lr 2.500000e-06	reg	1.000000e+01	batch	250	iters	1000	train accuracy: 0.318633	val accuracy: 0.335000
lr 2.500000e-06	reg	1.000000e+01	batch	250	iters	2000	train accuracy: 0.367469	val accuracy: 0.375000
lr 2.500000e-06	reg	1.000000e+01	batch	500	iters	250	train accuracy: 0.196224	val accuracy: 0.216000
lr 2.500000e-06	reg	1.000000e+01	batch	500	iters	500	train accuracy: 0.261122	val accuracy: 0.277000
lr 2.500000e-06	reg	1.000000e+01	batch	500	iters	1000	train accuracy: 0.327857	val accuracy: 0.332000
lr 2.500000e-06	reg	1.000000e+01	batch	500	iters	2000	train accuracy: 0.373469	val accuracy: 0.359000
lr 2.500000e-06	reg	1.000000e+01	batch	750	iters	250	train accuracy: 0.169490	val accuracy: 0.186000
lr 2.500000e-06	reg	1.000000e+01	batch	750	iters	500	train accuracy: 0.234694	val accuracy: 0.261000
lr 2.500000e-06	reg	1.000000e+01	batch	750	iters	1000	train accuracy: 0.327286	val accuracy: 0.341000
lr 2.500000e-06	reg	1.000000e+01	batch	750	iters	2000	train accuracy: 0.364571	val accuracy: 0.356000
lr 2.500000e-06	reg	2.500000e+01	batch	250	iters	250	train accuracy: 0.185633	val accuracy: 0.192000
lr 2.500000e-06	reg	2.500000e+01	batch	250	iters	500	train accuracy: 0.235816	val accuracy: 0.245000
lr 2.500000e-06	reg	2.500000e+01	batch	250	iters	1000	train accuracy: 0.311408	val accuracy: 0.305000
lr 2.500000e-06	reg	2.500000e+01	batch	250	iters	2000	train accuracy: 0.373490	val accuracy: 0.364000
lr 2.500000e-06	reg	2.500000e+01	batch	500	iters	250	train accuracy: 0.173837	val accuracy: 0.168000
lr 2.500000e-06	reg	2.500000e+01	batch	500	iters	500	train accuracy: 0.236082	val accuracy: 0.224000
lr 2.500000e-06	reg	2.500000e+01	batch	500	iters	1000	train accuracy: 0.318408	val accuracy: 0.304000
lr 2.500000e-06	reg	2.500000e+01	batch	500	iters	2000	train accuracy: 0.377184	val accuracy: 0.368000
lr 2.500000e-06	reg	2.500000e+01	batch	750	iters	250	train accuracy: 0.176633	val accuracy: 0.176000
lr 2.500000e-06	reg	2.500000e+01	batch	750	iters	500	train accuracy: 0.221163	val accuracy: 0.219000
lr 2.500000e-06	reg	2.500000e+01	batch	750	iters	1000	train accuracy: 0.320306	val accuracy: 0.322000
lr 2.500000e-06	reg	2.500000e+01	batch	750	iters	2000	train accuracy: 0.376755	val accuracy: 0.389000
lr 2.500000e-06	reg	5.000000e+01</						

lr 2.500000e-06	reg	2.500000e+04	batch	250	iters	2000	train accuracy: 0.413306	val accuracy: 0.413000
lr 2.500000e-06	reg	2.500000e+04	batch	500	iters	250	train accuracy: 0.411714	val accuracy: 0.410000
lr 2.500000e-06	reg	2.500000e+04	batch	500	iters	500	train accuracy: 0.417041	val accuracy: 0.418000
lr 2.500000e-06	reg	2.500000e+04	batch	500	iters	1000	train accuracy: 0.411531	val accuracy: 0.402000
lr 2.500000e-06	reg	2.500000e+04	batch	500	iters	2000	train accuracy: 0.417184	val accuracy: 0.414000
lr 2.500000e-06	reg	2.500000e+04	batch	750	iters	250	train accuracy: 0.413102	val accuracy: 0.411000
lr 2.500000e-06	reg	2.500000e+04	batch	750	iters	500	train accuracy: 0.410633	val accuracy: 0.396000
lr 2.500000e-06	reg	2.500000e+04	batch	750	iters	1000	train accuracy: 0.410612	val accuracy: 0.421000
lr 2.500000e-06	reg	2.500000e+04	batch	750	iters	2000	train accuracy: 0.416796	val accuracy: 0.421000
lr 2.500000e-06	reg	5.000000e+04	batch	250	iters	250	train accuracy: 0.392918	val accuracy: 0.403000
lr 2.500000e-06	reg	5.000000e+04	batch	250	iters	500	train accuracy: 0.401980	val accuracy: 0.404000
lr 2.500000e-06	reg	5.000000e+04	batch	250	iters	1000	train accuracy: 0.401857	val accuracy: 0.402000
lr 2.500000e-06	reg	5.000000e+04	batch	250	iters	2000	train accuracy: 0.387204	val accuracy: 0.378000
lr 2.500000e-06	reg	5.000000e+04	batch	500	iters	250	train accuracy: 0.408408	val accuracy: 0.407000
lr 2.500000e-06	reg	5.000000e+04	batch	500	iters	500	train accuracy: 0.397714	val accuracy: 0.408000
lr 2.500000e-06	reg	5.000000e+04	batch	500	iters	1000	train accuracy: 0.407735	val accuracy: 0.393000
lr 2.500000e-06	reg	5.000000e+04	batch	500	iters	2000	train accuracy: 0.403061	val accuracy: 0.397000
lr 2.500000e-06	reg	5.000000e+04	batch	750	iters	250	train accuracy: 0.412592	val accuracy: 0.423000
lr 2.500000e-06	reg	5.000000e+04	batch	750	iters	500	train accuracy: 0.411061	val accuracy: 0.412000
lr 2.500000e-06	reg	5.000000e+04	batch	750	iters	1000	train accuracy: 0.407122	val accuracy: 0.408000
lr 2.500000e-06	reg	5.000000e+04	batch	750	iters	2000	train accuracy: 0.408612	val accuracy: 0.423000
lr 2.500000e-06	reg	1.000000e+05	batch	250	iters	250	train accuracy: 0.378245	val accuracy: 0.349000
lr 2.500000e-06	reg	1.000000e+05	batch	250	iters	500	train accuracy: 0.390367	val accuracy: 0.397000
lr 2.500000e-06	reg	1.000000e+05	batch	250	iters	1000	train accuracy: 0.375898	val accuracy: 0.361000
lr 2.500000e-06	reg	1.000000e+05	batch	250	iters	2000	train accuracy: 0.392531	val accuracy: 0.377000
lr 2.500000e-06	reg	1.000000e+05	batch	500	iters	250	train accuracy: 0.401918	val accuracy: 0.375000
lr 2.500000e-06	reg	1.000000e+05	batch	500	iters	500	train accuracy: 0.411020	val accuracy: 0.417000
lr 2.500000e-06	reg	1.000000e+05	batch	500	iters	1000	train accuracy: 0.405531	val accuracy: 0.406000
lr 2.500000e-06	reg	1.000000e+05	batch	500	iters	2000	train accuracy: 0.396714	val accuracy: 0.372000
lr 2.500000e-06	reg	1.000000e+05	batch	750	iters	250	train accuracy: 0.401367	val accuracy: 0.408000
lr 2.500000e-06	reg	1.000000e+05	batch	750	iters	500	train accuracy: 0.407408	val accuracy: 0.393000
lr 2.500000e-06	reg	1.000000e+05	batch	750	iters	1000	train accuracy: 0.408204	val accuracy: 0.398000
lr 2.500000e-06	reg	1.000000e+05	batch	750	iters	2000	train accuracy: 0.394122	val accuracy: 0.394000
lr 2.500000e-06	reg	2.500000e+05	batch	250	iters	250	train accuracy: 0.314571	val accuracy: 0.305000
lr 2.500000e-06	reg	2.500000e+05	batch	250	iters	500	train accuracy: 0.299388	val accuracy: 0.293000
lr 2.500000e-06	reg	2.500000e+05	batch	250	iters	1000	train accuracy: 0.301122	val accuracy: 0.301000
lr 2.500000e-06	reg	2.500000e+05	batch	250	iters	2000	train accuracy: 0.302857	val accuracy: 0.317000
lr 2.500000e-06	reg	2.500000e+05</						

lr	2.500000e-06	reg	1.000000e+06	batch	750	iters	2000	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	2.500000e+06	batch	250	iters	250	train accuracy:	0.105959	val accuracy:	0.118000
lr	2.500000e-06	reg	2.500000e+06	batch	250	iters	500	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	2.500000e+06	batch	250	iters	1000	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	2.500000e+06	batch	250	iters	2000	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	2.500000e+06	batch	500	iters	250	train accuracy:	0.084082	val accuracy:	0.070000
lr	2.500000e-06	reg	2.500000e+06	batch	500	iters	500	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	2.500000e+06	batch	500	iters	1000	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	2.500000e+06	batch	500	iters	2000	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	2.500000e+06	batch	750	iters	250	train accuracy:	0.087571	val accuracy:	0.078000
lr	2.500000e-06	reg	2.500000e+06	batch	750	iters	500	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	2.500000e+06	batch	750	iters	1000	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	2.500000e+06	batch	750	iters	2000	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	5.000000e+06	batch	250	iters	250	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	5.000000e+06	batch	250	iters	500	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	5.000000e+06	batch	250	iters	1000	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	5.000000e+06	batch	250	iters	2000	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	5.000000e+06	batch	500	iters	250	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	5.000000e+06	batch	500	iters	500	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	5.000000e+06	batch	500	iters	1000	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	5.000000e+06	batch	500	iters	2000	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	5.000000e+06	batch	750	iters	250	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	5.000000e+06	batch	750	iters	500	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	5.000000e+06	batch	750	iters	1000	train accuracy:	0.100265	val accuracy:	0.087000
lr	2.500000e-06	reg	5.000000e+06	batch	750	iters	2000	train accuracy:	0.100265	val accuracy:	0.087000
lr	5.000000e-06	reg	1.000000e+01	batch	250	iters	250	train accuracy:	0.248633	val accuracy:	0.214000
lr	5.000000e-06	reg	1.000000e+01	batch	250	iters	500	train accuracy:	0.309184	val accuracy:	0.277000
lr	5.000000e-06	reg	1.000000e+01	batch	250	iters	1000	train accuracy:	0.364347	val accuracy:	0.362000
lr	5.000000e-06	reg	1.000000e+01	batch	250	iters	2000	train accuracy:	0.400265	val accuracy:	0.388000
lr	5.000000e-06	reg	1.000000e+01	batch	500	iters	250	train accuracy:	0.210061	val accuracy:	0.205000
lr	5.000000e-06	reg	1.000000e+01	batch	500	iters	500	train accuracy:	0.309245	val accuracy:	0.293000
lr	5.000000e-06	reg	1.000000e+01	batch	500	iters	1000	train accuracy:	0.379980	val accuracy:	0.378000
lr	5.000000e-06	reg	1.000000e+01	batch	500	iters	2000	train accuracy:	0.405551	val accuracy:	0.394000
lr	5.000000e-06	reg	1.000000e+01	batch	750	iters	250	train accuracy:	0.236000	val accuracy:	0.229000
lr	5.000000e-06	reg									

1r 5.000000e-06	reg	1.000000e+02	batch	500	iters	2000	train accuracy: 0.414020	val accuracy: 0.419000
1r 5.000000e-06	reg	1.000000e+02	batch	750	iters	250	train accuracy: 0.239429	val accuracy: 0.240000
1r 5.000000e-06	reg	1.000000e+02	batch	750	iters	500	train accuracy: 0.330163	val accuracy: 0.327000
1r 5.000000e-06	reg	1.000000e+02	batch	750	iters	1000	train accuracy: 0.403163	val accuracy: 0.401000
1r 5.000000e-06	reg	1.000000e+02	batch	750	iters	2000	train accuracy: 0.414082	val accuracy: 0.412000
1r 5.000000e-06	reg	2.500000e+02	batch	250	iters	250	train accuracy: 0.266224	val accuracy: 0.269000
1r 5.000000e-06	reg	2.500000e+02	batch	250	iters	500	train accuracy: 0.377939	val accuracy: 0.380000
1r 5.000000e-06	reg	2.500000e+02	batch	250	iters	1000	train accuracy: 0.412449	val accuracy: 0.412000
1r 5.000000e-06	reg	2.500000e+02	batch	250	iters	2000	train accuracy: 0.414347	val accuracy: 0.425000
1r 5.000000e-06	reg	2.500000e+02	batch	500	iters	250	train accuracy: 0.276388	val accuracy: 0.274000
1r 5.000000e-06	reg	2.500000e+02	batch	500	iters	500	train accuracy: 0.384653	val accuracy: 0.363000
1r 5.000000e-06	reg	2.500000e+02	batch	500	iters	1000	train accuracy: 0.412143	val accuracy: 0.408000
1r 5.000000e-06	reg	2.500000e+02	batch	500	iters	2000	train accuracy: 0.415469	val accuracy: 0.420000
1r 5.000000e-06	reg	2.500000e+02	batch	750	iters	250	train accuracy: 0.282388	val accuracy: 0.272000
1r 5.000000e-06	reg	2.500000e+02	batch	750	iters	500	train accuracy: 0.373163	val accuracy: 0.368000
1r 5.000000e-06	reg	2.500000e+02	batch	750	iters	1000	train accuracy: 0.415816	val accuracy: 0.424000
1r 5.000000e-06	reg	2.500000e+02	batch	750	iters	2000	train accuracy: 0.415714	val accuracy: 0.415000
1r 5.000000e-06	reg	5.000000e+02	batch	250	iters	250	train accuracy: 0.301592	val accuracy: 0.299000
1r 5.000000e-06	reg	5.000000e+02	batch	250	iters	500	train accuracy: 0.400959	val accuracy: 0.397000
1r 5.000000e-06	reg	5.000000e+02	batch	250	iters	1000	train accuracy: 0.412061	val accuracy: 0.417000
1r 5.000000e-06	reg	5.000000e+02	batch	250	iters	2000	train accuracy: 0.415224	val accuracy: 0.421000
1r 5.000000e-06	reg	5.000000e+02	batch	500	iters	250	train accuracy: 0.306408	val accuracy: 0.305000
1r 5.000000e-06	reg	5.000000e+02	batch	500	iters	500	train accuracy: 0.411653	val accuracy: 0.411000
1r 5.000000e-06	reg	5.000000e+02	batch	500	iters	1000	train accuracy: 0.414490	val accuracy: 0.420000
1r 5.000000e-06	reg	5.000000e+02	batch	500	iters	2000	train accuracy: 0.416082	val accuracy: 0.419000
1r 5.000000e-06	reg	5.000000e+02	batch	750	iters	250	train accuracy: 0.313347	val accuracy: 0.297000
1r 5.000000e-06	reg	5.000000e+02	batch	750	iters	500	train accuracy: 0.404633	val accuracy: 0.412000
1r 5.000000e-06	reg	5.000000e+02	batch	750	iters	1000	train accuracy: 0.415163	val accuracy: 0.424000
1r 5.000000e-06	reg	5.000000e+02	batch	750	iters	2000	train accuracy: 0.415673	val accuracy: 0.417000
1r 5.000000e-06	reg	1.000000e+03	batch	250	iters	250	train accuracy: 0.375878	val accuracy: 0.363000
1r 5.000000e-06	reg	1.000000e+03	batch	250	iters	500	train accuracy: 0.41612	val accuracy: 0.427000
1r 5.000000e-06	reg	1.000000e+03	batch	250	iters	1000	train accuracy: 0.414367	val accuracy: 0.418000
1r 5.000000e-06	reg	1.000000e+03	batch	250	iters	2000	train accuracy: 0.414816	val accuracy: 0.415000
1r 5.000000e-06	reg	1.000000e+03	batch	500	iters	250	train accuracy: 0.371143	val accuracy: 0.356000
1r 5.000000e-06	reg	1.000000e+03	batch	500	iters	500	train accuracy: 0.412592	val accuracy: 0.408000
1r 5.000000e-06	reg	1.000000e+03	batch	500	iters	1000	train accuracy: 0.414755	val accuracy: 0.413000
1r 5.000000e-06	reg	1.000000e+03	batch	500	iters	2000	train accuracy: 0.416041	val accuracy: 0.416000
1r								

1r 5.000000e-06	reg	1.000000e+04	batch	250	iters	2000	train accuracy: 0.409673	val accuracy: 0.432000
1r 5.000000e-06	reg	1.000000e+04	batch	500	iters	250	train accuracy: 0.416469	val accuracy: 0.415000
1r 5.000000e-06	reg	1.000000e+04	batch	500	iters	500	train accuracy: 0.409184	val accuracy: 0.401000
1r 5.000000e-06	reg	1.000000e+04	batch	500	iters	1000	train accuracy: 0.416306	val accuracy: 0.410000
1r 5.000000e-06	reg	1.000000e+04	batch	500	iters	2000	train accuracy: 0.412837	val accuracy: 0.415000
1r 5.000000e-06	reg	1.000000e+04	batch	750	iters	250	train accuracy: 0.414061	val accuracy: 0.420000
1r 5.000000e-06	reg	1.000000e+04	batch	750	iters	500	train accuracy: 0.414102	val accuracy: 0.407000
1r 5.000000e-06	reg	1.000000e+04	batch	750	iters	1000	train accuracy: 0.417673	val accuracy: 0.419000
1r 5.000000e-06	reg	1.000000e+04	batch	750	iters	2000	train accuracy: 0.412633	val accuracy: 0.412000
1r 5.000000e-06	reg	2.500000e+04	batch	250	iters	250	train accuracy: 0.404898	val accuracy: 0.381000
1r 5.000000e-06	reg	2.500000e+04	batch	250	iters	500	train accuracy: 0.398735	val accuracy: 0.400000
1r 5.000000e-06	reg	2.500000e+04	batch	250	iters	1000	train accuracy: 0.398939	val accuracy: 0.388000
1r 5.000000e-06	reg	2.500000e+04	batch	250	iters	2000	train accuracy: 0.388980	val accuracy: 0.398000
1r 5.000000e-06	reg	2.500000e+04	batch	500	iters	250	train accuracy: 0.411673	val accuracy: 0.413000
1r 5.000000e-06	reg	2.500000e+04	batch	500	iters	500	train accuracy: 0.402755	val accuracy: 0.384000
1r 5.000000e-06	reg	2.500000e+04	batch	500	iters	1000	train accuracy: 0.410102	val accuracy: 0.412000
1r 5.000000e-06	reg	2.500000e+04	batch	500	iters	2000	train accuracy: 0.406735	val accuracy: 0.401000
1r 5.000000e-06	reg	2.500000e+04	batch	750	iters	250	train accuracy: 0.413735	val accuracy: 0.413000
1r 5.000000e-06	reg	2.500000e+04	batch	750	iters	500	train accuracy: 0.416408	val accuracy: 0.412000
1r 5.000000e-06	reg	2.500000e+04	batch	750	iters	1000	train accuracy: 0.412265	val accuracy: 0.407000
1r 5.000000e-06	reg	2.500000e+04	batch	750	iters	2000	train accuracy: 0.406612	val accuracy: 0.399000
1r 5.000000e-06	reg	5.000000e+04	batch	250	iters	250	train accuracy: 0.379082	val accuracy: 0.371000
1r 5.000000e-06	reg	5.000000e+04	batch	250	iters	500	train accuracy: 0.376551	val accuracy: 0.357000
1r 5.000000e-06	reg	5.000000e+04	batch	250	iters	1000	train accuracy: 0.385245	val accuracy: 0.383000
1r 5.000000e-06	reg	5.000000e+04	batch	250	iters	2000	train accuracy: 0.380408	val accuracy: 0.372000
1r 5.000000e-06	reg	5.000000e+04	batch	500	iters	250	train accuracy: 0.397184	val accuracy: 0.397000
1r 5.000000e-06	reg	5.000000e+04	batch	500	iters	500	train accuracy: 0.388959	val accuracy: 0.381000
1r 5.000000e-06	reg	5.000000e+04	batch	500	iters	1000	train accuracy: 0.397959	val accuracy: 0.406000
1r 5.000000e-06	reg	5.000000e+04	batch	500	iters	2000	train accuracy: 0.402469	val accuracy: 0.407000
1r 5.000000e-06	reg	5.000000e+04	batch	750	iters	250	train accuracy: 0.401816	val accuracy: 0.386000
1r 5.000000e-06	reg	5.000000e+04	batch	750	iters	500	train accuracy: 0.407959	val accuracy: 0.407000
1r 5.000000e-06	reg	5.000000e+04	batch	750	iters	1000	train accuracy: 0.408367	val accuracy: 0.403000
1r 5.000000e-06	reg	5.000000e+04	batch	750	iters	2000	train accuracy: 0.400245	val accuracy: 0.391000
1r 5.000000e-06	reg	1.000000e+05	batch	250	iters	250	train accuracy: 0.324694	val accuracy: 0.309000
1r 5.000000e-06	reg	1.000000e+05	batch	250	iters	500	train accuracy: 0.334184	val accuracy: 0.348000
1r 5.000000e-06	reg	1.000000e+05	batch	250	iters	1000	train accuracy: 0.339878	val accuracy: 0.341000
1r 5.000000e-06	reg	1.000000e+05	batch	250	iters	2000	train accuracy: 0.348878	val accuracy: 0.368000
1								

[illegible]

lr	1.000000e-05	reg	5.000000e+01	batch	500	iters	2000	train accuracy: 0.414633	val accuracy: 0.412000
lr	1.000000e-05	reg	5.000000e+01	batch	750	iters	250	train accuracy: 0.324592	val accuracy: 0.342000
lr	1.000000e-05	reg	5.000000e+01	batch	750	iters	500	train accuracy: 0.381714	val accuracy: 0.347000
lr	1.000000e-05	reg	5.000000e+01	batch	750	iters	1000	train accuracy: 0.412204	val accuracy: 0.414000
lr	1.000000e-05	reg	5.000000e+01	batch	750	iters	2000	train accuracy: 0.415918	val accuracy: 0.415000
lr	1.000000e-05	reg	1.000000e+02	batch	250	iters	250	train accuracy: 0.336796	val accuracy: 0.332000
lr	1.000000e-05	reg	1.000000e+02	batch	250	iters	500	train accuracy: 0.409449	val accuracy: 0.412000
lr	1.000000e-05	reg	1.000000e+02	batch	250	iters	1000	train accuracy: 0.416592	val accuracy: 0.417000
lr	1.000000e-05	reg	1.000000e+02	batch	250	iters	2000	train accuracy: 0.414388	val accuracy: 0.420000
lr	1.000000e-05	reg	1.000000e+02	batch	500	iters	250	train accuracy: 0.335735	val accuracy: 0.330000
lr	1.000000e-05	reg	1.000000e+02	batch	500	iters	500	train accuracy: 0.390735	val accuracy: 0.374000
lr	1.000000e-05	reg	1.000000e+02	batch	500	iters	1000	train accuracy: 0.412633	val accuracy: 0.415000
lr	1.000000e-05	reg	1.000000e+02	batch	500	iters	2000	train accuracy: 0.415653	val accuracy: 0.418000
lr	1.000000e-05	reg	1.000000e+02	batch	750	iters	250	train accuracy: 0.336714	val accuracy: 0.342000
lr	1.000000e-05	reg	1.000000e+02	batch	750	iters	500	train accuracy: 0.398796	val accuracy: 0.411000
lr	1.000000e-05	reg	1.000000e+02	batch	750	iters	1000	train accuracy: 0.413571	val accuracy: 0.423000
lr	1.000000e-05	reg	1.000000e+02	batch	750	iters	2000	train accuracy: 0.415490	val accuracy: 0.424000
lr	1.000000e-05	reg	2.500000e+02	batch	250	iters	250	train accuracy: 0.371122	val accuracy: 0.362000
lr	1.000000e-05	reg	2.500000e+02	batch	250	iters	500	train accuracy: 0.412306	val accuracy: 0.424000
lr	1.000000e-05	reg	2.500000e+02	batch	250	iters	1000	train accuracy: 0.416673	val accuracy: 0.424000
lr	1.000000e-05	reg	2.500000e+02	batch	250	iters	2000	train accuracy: 0.416122	val accuracy: 0.419000
lr	1.000000e-05	reg	2.500000e+02	batch	500	iters	250	train accuracy: 0.370959	val accuracy: 0.379000
lr	1.000000e-05	reg	2.500000e+02	batch	500	iters	500	train accuracy: 0.411204	val accuracy: 0.420000
lr	1.000000e-05	reg	2.500000e+02	batch	500	iters	1000	train accuracy: 0.415204	val accuracy: 0.418000
lr	1.000000e-05	reg	2.500000e+02	batch	500	iters	2000	train accuracy: 0.416306	val accuracy: 0.416000
lr	1.000000e-05	reg	2.500000e+02	batch	750	iters	250	train accuracy: 0.364102	val accuracy: 0.363000
lr	1.000000e-05	reg	2.500000e+02	batch	750	iters	500	train accuracy: 0.413000	val accuracy: 0.416000
lr	1.000000e-05	reg	2.500000e+02	batch	750	iters	1000	train accuracy: 0.415714	val accuracy: 0.416000
lr	1.000000e-05	reg	2.500000e+02	batch	750	iters	2000	train accuracy: 0.414980	val accuracy: 0.420000
lr	1.000000e-05	reg	5.000000e+02	batch	250	iters	250	train accuracy: 0.403265	val accuracy: 0.407000
lr	1.000000e-05	reg	5.000000e+02	batch	250	iters	500	train accuracy: 0.413469	val accuracy: 0.409000
lr	1.000000e-05	reg	5.000000e+02	batch	250	iters	1000	train accuracy: 0.416898	val accuracy: 0.423000
lr	1.000000e-05	reg	5.000000e+02	batch	250	iters	2000	train accuracy: 0.414408	val accuracy: 0.414000
lr	1.000000e-05	reg	5.000000e+02	batch	500	iters	250	train accuracy: 0.400510	val accuracy: 0.392000
lr	1.000000e-05	reg	5.000000e+02	batch	500	iters	500	train accuracy: 0.414204	val accuracy: 0.409000
lr	1.000000e-05	reg	5.000000e+02	batch	500	iters	1000	train accuracy: 0.416551	val accuracy: 0.414000
lr									

lr 1.000000e-05 reg 5.000000e+03 batch 250 iters 2000 train accuracy: 0.405143 val accuracy: 0.383000
 lr 1.000000e-05 reg 5.000000e+03 batch 500 iters 250 train accuracy: 0.405510 val accuracy: 0.396000
 lr 1.000000e-05 reg 5.000000e+03 batch 500 iters 500 train accuracy: 0.409776 val accuracy: 0.412000
 lr 1.000000e-05 reg 5.000000e+03 batch 500 iters 1000 train accuracy: 0.405980 val accuracy: 0.402000
 lr 1.000000e-05 reg 5.000000e+03 batch 500 iters 2000 train accuracy: 0.405163 val accuracy: 0.397000
 lr 1.000000e-05 reg 5.000000e+03 batch 750 iters 250 train accuracy: 0.408122 val accuracy: 0.405000
 lr 1.000000e-05 reg 5.000000e+03 batch 750 iters 500 train accuracy: 0.417327 val accuracy: 0.422000
 lr 1.000000e-05 reg 5.000000e+03 batch 750 iters 1000 train accuracy: 0.415796 val accuracy: 0.413000
 lr 1.000000e-05 reg 5.000000e+03 batch 750 iters 2000 train accuracy: 0.415388 val accuracy: 0.424000
 lr 1.000000e-05 reg 1.000000e+04 batch 250 iters 250 train accuracy: 0.406061 val accuracy: 0.411000
 lr 1.000000e-05 reg 1.000000e+04 batch 250 iters 500 train accuracy: 0.408531 val accuracy: 0.411000
 lr 1.000000e-05 reg 1.000000e+04 batch 250 iters 1000 train accuracy: 0.403959 val accuracy: 0.407000
 lr 1.000000e-05 reg 1.000000e+04 batch 250 iters 2000 train accuracy: 0.386061 val accuracy: 0.368000
 lr 1.000000e-05 reg 1.000000e+04 batch 500 iters 250 train accuracy: 0.410918 val accuracy: 0.415000
 lr 1.000000e-05 reg 1.000000e+04 batch 500 iters 500 train accuracy: 0.409102 val accuracy: 0.415000
 lr 1.000000e-05 reg 1.000000e+04 batch 500 iters 1000 train accuracy: 0.408959 val accuracy: 0.412000
 lr 1.000000e-05 reg 1.000000e+04 batch 500 iters 2000 train accuracy: 0.409571 val accuracy: 0.423000
 lr 1.000000e-05 reg 1.000000e+04 batch 750 iters 250 train accuracy: 0.414286 val accuracy: 0.407000
 lr 1.000000e-05 reg 1.000000e+04 batch 750 iters 500 train accuracy: 0.404082 val accuracy: 0.402000
 lr 1.000000e-05 reg 1.000000e+04 batch 750 iters 1000 train accuracy: 0.414102 val accuracy: 0.406000
 lr 1.000000e-05 reg 1.000000e+04 batch 750 iters 2000 train accuracy: 0.406735 val accuracy: 0.399000
 lr 1.000000e-05 reg 2.500000e+04 batch 250 iters 250 train accuracy: 0.368918 val accuracy: 0.349000
 lr 1.000000e-05 reg 2.500000e+04 batch 250 iters 500 train accuracy: 0.365082 val accuracy: 0.339000
 lr 1.000000e-05 reg 2.500000e+04 batch 250 iters 1000 train accuracy: 0.384204 val accuracy: 0.367000
 lr 1.000000e-05 reg 2.500000e+04 batch 250 iters 2000 train accuracy: 0.389082 val accuracy: 0.388000
 lr 1.000000e-05 reg 2.500000e+04 batch 500 iters 250 train accuracy: 0.384531 val accuracy: 0.396000
 lr 1.000000e-05 reg 2.500000e+04 batch 500 iters 500 train accuracy: 0.396551 val accuracy: 0.413000
 lr 1.000000e-05 reg 2.500000e+04 batch 500 iters 1000 train accuracy: 0.379102 val accuracy: 0.378000
 lr 1.000000e-05 reg 2.500000e+04 batch 500 iters 2000 train accuracy: 0.387163 val accuracy: 0.375000
 lr 1.000000e-05 reg 2.500000e+04 batch 750 iters 250 train accuracy: 0.403122 val accuracy: 0.395000
 lr 1.000000e-05 reg 2.500000e+04 batch 750 iters 500 train accuracy: 0.404041 val accuracy: 0.400000
 lr 1.000000e-05 reg 2.500000e+04 batch 750 iters 1000 train accuracy: 0.399020 val accuracy: 0.407000
 lr 1.000000e-05 reg 2.500000e+04 batch 750 iters 2000 train accuracy: 0.395286 val accuracy: 0.396000
 lr 1.000000e-05 reg 5.000000e+04 batch 250 iters 250 train accuracy: 0.330939 val accuracy: 0.349000
 lr 1.000000e-05 reg 5.000000e+04 batch 250 iters 500 train accuracy: 0.346551 val accuracy: 0.319000
 lr 1.000000e-05 reg 5.000000e+04 batch 250 iters 1000 train accuracy: 0.344490 val accuracy: 0.336000
 lr 1.000000e-05 reg 5.000000e+04 batch 250 iters 2000 train accuracy: 0.342796 val accuracy: 0.323000
 lr 1.000000e-05 reg 5.000000e+04 batch 500 iters 250 train accuracy: 0.343102 val accuracy: 0.335000
 lr 1.000000e-05 reg 5.000000e+04 batch 500 iters 500 train accuracy: 0.362469 val accuracy: 0.330000
 lr 1.000000e-05 reg 5.000000e+04 batch 500 iters 1000 train accuracy: 0.343755 val accuracy: 0.317000
 lr 1.000000e-05 reg 5.000000e+04 batch 500 iters 2000 train accuracy: 0.365388 val accuracy: 0.356000
 lr 1.000000e-05 reg 5.000000e+04 batch 750 iters 250 train accuracy: 0.371918 val accuracy: 0.386000
 lr 1.000000e-05 reg 5.000000e+04 batch 750 iters 500 train accuracy: 0.384816 val accuracy: 0.371000
 lr 1.000000e-05 reg 5.000000e+04 batch 750 iters 1000 train accuracy: 0.384143 val accuracy: 0.394000
 lr 1.000000e-05 reg 5.000000e+04 batch 750 iters 2000 train accuracy: 0.377286 val accuracy: 0.387000
 lr 1.000000e-05 reg 1.000000e+05 batch 250 iters 250 train accuracy: 0.105367 val accuracy: 0.087000
 lr 1.000000e-05 reg 1.000000e+05 batch 250 iters 500 train accuracy: 0.100102 val accuracy: 0.101000
 lr 1.000000e-05 reg 1.000000e+05 batch 250 iters 1000 train accuracy: 0.094714 val accuracy: 0.085000
 lr 1.000000e-05 reg 1.000000e+05 batch 250 iters 2000 train accuracy: 0.096286 val accuracy: 0.094000
 lr 1.000000e-05 reg 1.000000e+05 batch 500 iters 250 train accuracy: 0.097449 val accuracy: 0.105000
 lr 1.000000e-05 reg 1.000000e+05 batch 500 iters 500 train accuracy: 0.085918 val accuracy: 0.084000
 lr 1.000000e-05 reg 1.000000e+05 batch 500 iters 1000 train accuracy: 0.084510 val accuracy: 0.091000
 lr 1.000000e-05 reg 1.000000e+05 batch 500 iters 2000 train accuracy: 0.115490 val accuracy: 0.106000
 lr 1.000000e-05 reg 1.000000e+05 batch 750 iters 250 train accuracy: 0.093388 val accuracy: 0.105000
 lr 1.000000e-05 reg 1.000000e+05 batch 750 iters 500 train accuracy: 0.109020 val accuracy: 0.107000
 lr 1.000000e-05 reg 1.000000e+05 batch 750 iters 1000 train accuracy: 0.084571 val accuracy: 0.089000
 lr 1.000000e-05 reg 1.000000e+05 batch 750 iters 2000 train accuracy: 0.094796 val accuracy: 0.096000
 lr 1.000000e-05 reg 2.500000e+05 batch 250 iters 250 train accuracy: 0.104163 val accuracy: 0.105000
 lr 1.000000e-05 reg 2.500000e+05 batch 250 iters 500 train accuracy: 0.097020 val accuracy: 0.076000
 lr 1.000000e-05 reg 2.500000e+05 batch 250 iters 1000 train accuracy: 0.100265 val accuracy: 0.087000
 lr 1.000000e-05 reg 2.500000e+05 batch 250 iters 2000 train accuracy: 0.100265 val accuracy: 0.087000
 lr 1.000000e-05 reg 2.500000e+05 batch 500 iters 250 train accuracy: 0.091551 val accuracy: 0.086000
 lr 1.000000e-05 reg 2.500000e+05 batch 500 iters 500 train accuracy: 0.075102 val accuracy: 0.071000
 lr 1.000000e-05 reg 2.500000e+05 batch 500 iters 1000 train accuracy: 0.100265 val accuracy: 0.087000
 lr 1.000000e-05 reg 2.500000e+05 batch 500 iters 2000 train accuracy: 0.100265 val accuracy: 0.087000
 lr 1.000000e-05 reg 2.500000e+05 batch 750 iters 250 train accuracy: 0.086551 val accuracy: 0.100000
 lr 1.000000e-05 reg 2.500000e+05 batch 750 iters 500 train accuracy: 0.108327 val accuracy: 0.116000
 lr 1.000000e-05 reg 2.500000e+05 batch 750 iters 1000 train accuracy: 0.100265 val accuracy: 0.087000

[illegible]

lr 2.500000e-05	reg	2.500000e+01	batch	500	iters	2000	train accuracy: 0.416184	val accuracy: 0.420000
lr 2.500000e-05	reg	2.500000e+01	batch	750	iters	250	train accuracy: 0.386939	val accuracy: 0.385000
lr 2.500000e-05	reg	2.500000e+01	batch	750	iters	500	train accuracy: 0.411939	val accuracy: 0.413000
lr 2.500000e-05	reg	2.500000e+01	batch	750	iters	1000	train accuracy: 0.414673	val accuracy: 0.420000
lr 2.500000e-05	reg	2.500000e+01	batch	750	iters	2000	train accuracy: 0.415061	val accuracy: 0.415000
lr 2.500000e-05	reg	5.000000e+01	batch	250	iters	250	train accuracy: 0.399122	val accuracy: 0.396000
lr 2.500000e-05	reg	5.000000e+01	batch	250	iters	500	train accuracy: 0.414837	val accuracy: 0.415000
lr 2.500000e-05	reg	5.000000e+01	batch	250	iters	1000	train accuracy: 0.415857	val accuracy: 0.417000
lr 2.500000e-05	reg	5.000000e+01	batch	250	iters	2000	train accuracy: 0.416551	val accuracy: 0.416000
lr 2.500000e-05	reg	5.000000e+01	batch	500	iters	250	train accuracy: 0.400551	val accuracy: 0.394000
lr 2.500000e-05	reg	5.000000e+01	batch	500	iters	500	train accuracy: 0.411959	val accuracy: 0.408000
lr 2.500000e-05	reg	5.000000e+01	batch	500	iters	1000	train accuracy: 0.413490	val accuracy: 0.416000
lr 2.500000e-05	reg	5.000000e+01	batch	500	iters	2000	train accuracy: 0.414612	val accuracy: 0.421000
lr 2.500000e-05	reg	5.000000e+01	batch	750	iters	250	train accuracy: 0.401163	val accuracy: 0.396000
lr 2.500000e-05	reg	5.000000e+01	batch	750	iters	500	train accuracy: 0.414837	val accuracy: 0.416000
lr 2.500000e-05	reg	5.000000e+01	batch	750	iters	1000	train accuracy: 0.414633	val accuracy: 0.415000
lr 2.500000e-05	reg	5.000000e+01	batch	750	iters	2000	train accuracy: 0.415980	val accuracy: 0.417000
lr 2.500000e-05	reg	1.000000e+02	batch	250	iters	250	train accuracy: 0.413184	val accuracy: 0.408000
lr 2.500000e-05	reg	1.000000e+02	batch	250	iters	500	train accuracy: 0.414122	val accuracy: 0.407000
lr 2.500000e-05	reg	1.000000e+02	batch	250	iters	1000	train accuracy: 0.417082	val accuracy: 0.417000
lr 2.500000e-05	reg	1.000000e+02	batch	250	iters	2000	train accuracy: 0.416367	val accuracy: 0.425000
lr 2.500000e-05	reg	1.000000e+02	batch	500	iters	250	train accuracy: 0.408878	val accuracy: 0.402000
lr 2.500000e-05	reg	1.000000e+02	batch	500	iters	500	train accuracy: 0.414367	val accuracy: 0.411000
lr 2.500000e-05	reg	1.000000e+02	batch	500	iters	1000	train accuracy: 0.414612	val accuracy: 0.415000
lr 2.500000e-05	reg	1.000000e+02	batch	500	iters	2000	train accuracy: 0.414020	val accuracy: 0.417000
lr 2.500000e-05	reg	1.000000e+02	batch	750	iters	250	train accuracy: 0.408490	val accuracy: 0.408000
lr 2.500000e-05	reg	1.000000e+02	batch	750	iters	500	train accuracy: 0.414163	val accuracy: 0.418000
lr 2.500000e-05	reg	1.000000e+02	batch	750	iters	1000	train accuracy: 0.414224	val accuracy: 0.421000
lr 2.500000e-05	reg	1.000000e+02	batch	750	iters	2000	train accuracy: 0.416122	val accuracy: 0.420000
lr 2.500000e-05	reg	2.500000e+02	batch	250	iters	250	train accuracy: 0.411571	val accuracy: 0.423000
lr 2.500000e-05	reg	2.500000e+02	batch	250	iters	500	train accuracy: 0.412020	val accuracy: 0.410000
lr 2.500000e-05	reg	2.500000e+02	batch	250	iters	1000	train accuracy: 0.416224	val accuracy: 0.424000
lr 2.500000e-05	reg	2.500000e+02	batch	250	iters	2000	train accuracy: 0.416020	val accuracy: 0.414000
lr 2.500000e-05	reg	2.500000e+02	batch	500	iters	250	train accuracy: 0.414551	val accuracy: 0.416000
lr 2.500000e-05	reg	2.500000e+02	batch	500	iters	500	train accuracy: 0.416388	val accuracy: 0.422000
lr 2.500000e-05	reg	2.500000e+02	batch	500	iters	1000	train accuracy: 0.417000	val accuracy: 0.422000
lr 2.500000e-05	reg	2.500000e+02	batch	500	iters	2000	train accuracy: 0.416980	val accuracy: 0.419000
lr 2.500000e-05	reg	2.500000e+02</						

lr 2.500000e-05	reg	2.500000e+03	batch	250	iters	2000	train accuracy: 0.407673	val accuracy: 0.403000
lr 2.500000e-05	reg	2.500000e+03	batch	500	iters	250	train accuracy: 0.406388	val accuracy: 0.402000
lr 2.500000e-05	reg	2.500000e+03	batch	500	iters	500	train accuracy: 0.418551	val accuracy: 0.419000
lr 2.500000e-05	reg	2.500000e+03	batch	500	iters	1000	train accuracy: 0.414102	val accuracy: 0.401000
lr 2.500000e-05	reg	2.500000e+03	batch	500	iters	2000	train accuracy: 0.403143	val accuracy: 0.392000
lr 2.500000e-05	reg	2.500000e+03	batch	750	iters	250	train accuracy: 0.418490	val accuracy: 0.412000
lr 2.500000e-05	reg	2.500000e+03	batch	750	iters	500	train accuracy: 0.410857	val accuracy: 0.411000
lr 2.500000e-05	reg	2.500000e+03	batch	750	iters	1000	train accuracy: 0.416510	val accuracy: 0.418000
lr 2.500000e-05	reg	2.500000e+03	batch	750	iters	2000	train accuracy: 0.410388	val accuracy: 0.423000
lr 2.500000e-05	reg	5.000000e+03	batch	250	iters	250	train accuracy: 0.396306	val accuracy: 0.379000
lr 2.500000e-05	reg	5.000000e+03	batch	250	iters	500	train accuracy: 0.395184	val accuracy: 0.390000
lr 2.500000e-05	reg	5.000000e+03	batch	250	iters	1000	train accuracy: 0.400898	val accuracy: 0.387000
lr 2.500000e-05	reg	5.000000e+03	batch	250	iters	2000	train accuracy: 0.402041	val accuracy: 0.414000
lr 2.500000e-05	reg	5.000000e+03	batch	500	iters	250	train accuracy: 0.413469	val accuracy: 0.420000
lr 2.500000e-05	reg	5.000000e+03	batch	500	iters	500	train accuracy: 0.401918	val accuracy: 0.399000
lr 2.500000e-05	reg	5.000000e+03	batch	500	iters	1000	train accuracy: 0.393735	val accuracy: 0.392000
lr 2.500000e-05	reg	5.000000e+03	batch	500	iters	2000	train accuracy: 0.413367	val accuracy: 0.418000
lr 2.500000e-05	reg	5.000000e+03	batch	750	iters	250	train accuracy: 0.406531	val accuracy: 0.407000
lr 2.500000e-05	reg	5.000000e+03	batch	750	iters	500	train accuracy: 0.407102	val accuracy: 0.396000
lr 2.500000e-05	reg	5.000000e+03	batch	750	iters	1000	train accuracy: 0.404898	val accuracy: 0.397000
lr 2.500000e-05	reg	5.000000e+03	batch	750	iters	2000	train accuracy: 0.413224	val accuracy: 0.420000
lr 2.500000e-05	reg	1.000000e+04	batch	250	iters	250	train accuracy: 0.381878	val accuracy: 0.371000
lr 2.500000e-05	reg	1.000000e+04	batch	250	iters	500	train accuracy: 0.382020	val accuracy: 0.371000
lr 2.500000e-05	reg	1.000000e+04	batch	250	iters	1000	train accuracy: 0.359388	val accuracy: 0.350000
lr 2.500000e-05	reg	1.000000e+04	batch	250	iters	2000	train accuracy: 0.367224	val accuracy: 0.368000
lr 2.500000e-05	reg	1.000000e+04	batch	500	iters	250	train accuracy: 0.396306	val accuracy: 0.394000
lr 2.500000e-05	reg	1.000000e+04	batch	500	iters	500	train accuracy: 0.378020	val accuracy: 0.371000
lr 2.500000e-05	reg	1.000000e+04	batch	500	iters	1000	train accuracy: 0.398469	val accuracy: 0.405000
lr 2.500000e-05	reg	1.000000e+04	batch	500	iters	2000	train accuracy: 0.389694	val accuracy: 0.391000
lr 2.500000e-05	reg	1.000000e+04	batch	750	iters	250	train accuracy: 0.399102	val accuracy: 0.396000
lr 2.500000e-05	reg	1.000000e+04	batch	750	iters	500	train accuracy: 0.403735	val accuracy: 0.411000
lr 2.500000e-05	reg	1.000000e+04	batch	750	iters	1000	train accuracy: 0.397449	val accuracy: 0.404000
lr 2.500000e-05	reg	1.000000e+04	batch	750	iters	2000	train accuracy: 0.401388	val accuracy: 0.386000
lr 2.500000e-05	reg	2.500000e+04	batch	250	iters	250	train accuracy: 0.319673	val accuracy: 0.299000
lr 2.500000e-05	reg	2.500000e+04	batch	250	iters	500	train accuracy: 0.292367	val accuracy: 0.320000
lr 2.500000e-05	reg	2.500000e+04	batch	250	iters	1000	train accuracy: 0.319102	val accuracy: 0.288000
lr 2.500000e-05	reg	2.500000e+04	batch	250	iters	2000	train accuracy: 0.291347	val accuracy: 0.295000
lr 2.500000e-05	reg	2.500000e+04</						

[illegible]

1r	5.000000e-05	reg	1.000000e+01	batch	500	iters	2000	train accuracy: 0.424224	val accuracy: 0.428000
1r	5.000000e-05	reg	1.000000e+01	batch	750	iters	250	train accuracy: 0.406735	val accuracy: 0.409000
1r	5.000000e-05	reg	1.000000e+01	batch	750	iters	500	train accuracy: 0.411755	val accuracy: 0.412000
1r	5.000000e-05	reg	1.000000e+01	batch	750	iters	1000	train accuracy: 0.415857	val accuracy: 0.418000
1r	5.000000e-05	reg	1.000000e+01	batch	750	iters	2000	train accuracy: 0.424490	val accuracy: 0.432000
1r	5.000000e-05	reg	2.500000e+01	batch	250	iters	250	train accuracy: 0.411531	val accuracy: 0.404000
1r	5.000000e-05	reg	2.500000e+01	batch	250	iters	500	train accuracy: 0.412959	val accuracy: 0.417000
1r	5.000000e-05	reg	2.500000e+01	batch	250	iters	1000	train accuracy: 0.415367	val accuracy: 0.419000
1r	5.000000e-05	reg	2.500000e+01	batch	250	iters	2000	train accuracy: 0.415286	val accuracy: 0.416000
1r	5.000000e-05	reg	2.500000e+01	batch	500	iters	250	train accuracy: 0.413224	val accuracy: 0.408000
1r	5.000000e-05	reg	2.500000e+01	batch	500	iters	500	train accuracy: 0.413918	val accuracy: 0.422000
1r	5.000000e-05	reg	2.500000e+01	batch	500	iters	1000	train accuracy: 0.415490	val accuracy: 0.419000
1r	5.000000e-05	reg	2.500000e+01	batch	500	iters	2000	train accuracy: 0.414612	val accuracy: 0.420000
1r	5.000000e-05	reg	2.500000e+01	batch	750	iters	250	train accuracy: 0.409367	val accuracy: 0.414000
1r	5.000000e-05	reg	2.500000e+01	batch	750	iters	500	train accuracy: 0.413816	val accuracy: 0.418000
1r	5.000000e-05	reg	2.500000e+01	batch	750	iters	1000	train accuracy: 0.415408	val accuracy: 0.420000
1r	5.000000e-05	reg	2.500000e+01	batch	750	iters	2000	train accuracy: 0.415551	val accuracy: 0.415000
1r	5.000000e-05	reg	5.000000e+01	batch	250	iters	250	train accuracy: 0.415755	val accuracy: 0.421000
1r	5.000000e-05	reg	5.000000e+01	batch	250	iters	500	train accuracy: 0.415735	val accuracy: 0.420000
1r	5.000000e-05	reg	5.000000e+01	batch	250	iters	1000	train accuracy: 0.415694	val accuracy: 0.419000
1r	5.000000e-05	reg	5.000000e+01	batch	250	iters	2000	train accuracy: 0.413878	val accuracy: 0.417000
1r	5.000000e-05	reg	5.000000e+01	batch	500	iters	250	train accuracy: 0.415878	val accuracy: 0.416000
1r	5.000000e-05	reg	5.000000e+01	batch	500	iters	500	train accuracy: 0.415041	val accuracy: 0.416000
1r	5.000000e-05	reg	5.000000e+01	batch	500	iters	1000	train accuracy: 0.415367	val accuracy: 0.419000
1r	5.000000e-05	reg	5.000000e+01	batch	500	iters	2000	train accuracy: 0.415082	val accuracy: 0.422000
1r	5.000000e-05	reg	5.000000e+01	batch	750	iters	250	train accuracy: 0.411857	val accuracy: 0.420000
1r	5.000000e-05	reg	5.000000e+01	batch	750	iters	500	train accuracy: 0.413347	val accuracy: 0.411000
1r	5.000000e-05	reg	5.000000e+01	batch	750	iters	1000	train accuracy: 0.414898	val accuracy: 0.420000
1r	5.000000e-05	reg	5.000000e+01	batch	750	iters	2000	train accuracy: 0.415306	val accuracy: 0.419000
1r	5.000000e-05	reg	1.000000e+02	batch	250	iters	250	train accuracy: 0.417020	val accuracy: 0.413000
1r	5.000000e-05	reg	1.000000e+02	batch	250	iters	500	train accuracy: 0.414755	val accuracy: 0.421000
1r	5.000000e-05	reg	1.000000e+02	batch	250	iters	1000	train accuracy: 0.411755	val accuracy: 0.419000
1r	5.000000e-05	reg	1.000000e+02	batch	250	iters	2000	train accuracy: 0.415020	val accuracy: 0.409000
1r	5.000000e-05	reg	1.000000e+02	batch	500	iters	250	train accuracy: 0.413653	val accuracy: 0.416000
1r	5.000000e-05	reg	1.000000e+02	batch	500	iters	500	train accuracy: 0.413265	val accuracy: 0.406000
1r	5.000000e-05	reg	1.000000e+02	batch	500	iters	1000	train accuracy: 0.	

1r 5.000000e-05	reg	1.000000e+03	batch	250	iters	2000	train accuracy: 0.407673	val accuracy: 0.408000
1r 5.000000e-05	reg	1.000000e+03	batch	500	iters	250	train accuracy: 0.418551	val accuracy: 0.428000
1r 5.000000e-05	reg	1.000000e+03	batch	500	iters	500	train accuracy: 0.411959	val accuracy: 0.406000
1r 5.000000e-05	reg	1.000000e+03	batch	500	iters	1000	train accuracy: 0.413143	val accuracy: 0.417000
1r 5.000000e-05	reg	1.000000e+03	batch	500	iters	2000	train accuracy: 0.407673	val accuracy: 0.409000
1r 5.000000e-05	reg	1.000000e+03	batch	750	iters	250	train accuracy: 0.413816	val accuracy: 0.416000
1r 5.000000e-05	reg	1.000000e+03	batch	750	iters	500	train accuracy: 0.418347	val accuracy: 0.427000
1r 5.000000e-05	reg	1.000000e+03	batch	750	iters	1000	train accuracy: 0.416367	val accuracy: 0.415000
1r 5.000000e-05	reg	1.000000e+03	batch	750	iters	2000	train accuracy: 0.408429	val accuracy: 0.412000
1r 5.000000e-05	reg	2.500000e+03	batch	250	iters	250	train accuracy: 0.399796	val accuracy: 0.384000
1r 5.000000e-05	reg	2.500000e+03	batch	250	iters	500	train accuracy: 0.394653	val accuracy: 0.393000
1r 5.000000e-05	reg	2.500000e+03	batch	250	iters	1000	train accuracy: 0.393041	val accuracy: 0.412000
1r 5.000000e-05	reg	2.500000e+03	batch	250	iters	2000	train accuracy: 0.396673	val accuracy: 0.394000
1r 5.000000e-05	reg	2.500000e+03	batch	500	iters	250	train accuracy: 0.409347	val accuracy: 0.405000
1r 5.000000e-05	reg	2.500000e+03	batch	500	iters	500	train accuracy: 0.404816	val accuracy: 0.402000
1r 5.000000e-05	reg	2.500000e+03	batch	500	iters	1000	train accuracy: 0.407020	val accuracy: 0.410000
1r 5.000000e-05	reg	2.500000e+03	batch	500	iters	2000	train accuracy: 0.402837	val accuracy: 0.407000
1r 5.000000e-05	reg	2.500000e+03	batch	750	iters	250	train accuracy: 0.414531	val accuracy: 0.399000
1r 5.000000e-05	reg	2.500000e+03	batch	750	iters	500	train accuracy: 0.407449	val accuracy: 0.411000
1r 5.000000e-05	reg	2.500000e+03	batch	750	iters	1000	train accuracy: 0.407306	val accuracy: 0.386000
1r 5.000000e-05	reg	2.500000e+03	batch	750	iters	2000	train accuracy: 0.416102	val accuracy: 0.430000
1r 5.000000e-05	reg	5.000000e+03	batch	250	iters	250	train accuracy: 0.355898	val accuracy: 0.333000
1r 5.000000e-05	reg	5.000000e+03	batch	250	iters	500	train accuracy: 0.366347	val accuracy: 0.374000
1r 5.000000e-05	reg	5.000000e+03	batch	250	iters	1000	train accuracy: 0.380673	val accuracy: 0.409000
1r 5.000000e-05	reg	5.000000e+03	batch	250	iters	2000	train accuracy: 0.381347	val accuracy: 0.365000
1r 5.000000e-05	reg	5.000000e+03	batch	500	iters	250	train accuracy: 0.395653	val accuracy: 0.388000
1r 5.000000e-05	reg	5.000000e+03	batch	500	iters	500	train accuracy: 0.394306	val accuracy: 0.389000
1r 5.000000e-05	reg	5.000000e+03	batch	500	iters	1000	train accuracy: 0.403163	val accuracy: 0.404000
1r 5.000000e-05	reg	5.000000e+03	batch	500	iters	2000	train accuracy: 0.394653	val accuracy: 0.387000
1r 5.000000e-05	reg	5.000000e+03	batch	750	iters	250	train accuracy: 0.401857	val accuracy: 0.408000
1r 5.000000e-05	reg	5.000000e+03	batch	750	iters	500	train accuracy: 0.402122	val accuracy: 0.412000
1r 5.000000e-05	reg	5.000000e+03	batch	750	iters	1000	train accuracy: 0.417061	val accuracy: 0.435000
1r 5.000000e-05	reg	5.000000e+03	batch	750	iters	2000	train accuracy: 0.399122	val accuracy: 0.396000
1r 5.000000e-05	reg	1.000000e+04	batch	250	iters	250	train accuracy: 0.322408	val accuracy: 0.321000
1r 5.000000e-05	reg	1.000000e+04	batch	250	iters	500	train accuracy: 0.317388	val accuracy: 0.339000
1r 5.000000e-05	reg	1.000000e+04	batch	250	iters	1000	train accuracy: 0.333837	val accuracy: 0.334000
1r 5.000000e-05	reg	1.000000e+04	batch	250	iters	2000	train accuracy: 0.333612	val accuracy: 0.321000
1								

[illegible]

1r	5.000000e-05	reg	5.000000e+06	batch	500	iters	2000	train accuracy:	0.100265	val accuracy:	0.087000
1r	5.000000e-05	reg	5.000000e+06	batch	750	iters	250	train accuracy:	0.100265	val accuracy:	0.087000
1r	5.000000e-05	reg	5.000000e+06	batch	750	iters	500	train accuracy:	0.100265	val accuracy:	0.087000
1r	5.000000e-05	reg	5.000000e+06	batch	750	iters	1000	train accuracy:	0.100265	val accuracy:	0.087000
1r	5.000000e-05	reg	5.000000e+06	batch	750	iters	2000	train accuracy:	0.100265	val accuracy:	0.087000
1r	1.000000e-04	reg	1.000000e+01	batch	250	iters	250	train accuracy:	0.414469	val accuracy:	0.414000
1r	1.000000e-04	reg	1.000000e+01	batch	250	iters	500	train accuracy:	0.413980	val accuracy:	0.420000
1r	1.000000e-04	reg	1.000000e+01	batch	250	iters	1000	train accuracy:	0.424694	val accuracy:	0.426000
1r	1.000000e-04	reg	1.000000e+01	batch	250	iters	2000	train accuracy:	0.430347	val accuracy:	0.434000
1r	1.000000e-04	reg	1.000000e+01	batch	500	iters	250	train accuracy:	0.412102	val accuracy:	0.416000
1r	1.000000e-04	reg	1.000000e+01	batch	500	iters	500	train accuracy:	0.417388	val accuracy:	0.418000
1r	1.000000e-04	reg	1.000000e+01	batch	500	iters	1000	train accuracy:	0.423755	val accuracy:	0.428000
1r	1.000000e-04	reg	1.000000e+01	batch	500	iters	2000	train accuracy:	0.431041	val accuracy:	0.435000
1r	1.000000e-04	reg	1.000000e+01	batch	750	iters	250	train accuracy:	0.414388	val accuracy:	0.424000
1r	1.000000e-04	reg	1.000000e+01	batch	750	iters	500	train accuracy:	0.416735	val accuracy:	0.425000
1r	1.000000e-04	reg	1.000000e+01	batch	750	iters	1000	train accuracy:	0.424061	val accuracy:	0.431000
1r	1.000000e-04	reg	1.000000e+01	batch	750	iters	2000	train accuracy:	0.429571	val accuracy:	0.439000
1r	1.000000e-04	reg	2.500000e+01	batch	250	iters	250	train accuracy:	0.416673	val accuracy:	0.415000
1r	1.000000e-04	reg	2.500000e+01	batch	250	iters	500	train accuracy:	0.413143	val accuracy:	0.416000
1r	1.000000e-04	reg	2.500000e+01	batch	250	iters	1000	train accuracy:	0.413490	val accuracy:	0.419000
1r	1.000000e-04	reg	2.500000e+01	batch	250	iters	2000	train accuracy:	0.414490	val accuracy:	0.419000
1r	1.000000e-04	reg	2.500000e+01	batch	500	iters	250	train accuracy:	0.417163	val accuracy:	0.417000
1r	1.000000e-04	reg	2.500000e+01	batch	500	iters	500	train accuracy:	0.416571	val accuracy:	0.415000
1r	1.000000e-04	reg	2.500000e+01	batch	500	iters	1000	train accuracy:	0.413796	val accuracy:	0.418000
1r	1.000000e-04	reg	2.500000e+01	batch	500	iters	2000	train accuracy:	0.416082	val accuracy:	0.420000
1r	1.000000e-04	reg	2.500000e+01	batch	750	iters	250	train accuracy:	0.414592	val accuracy:	0.416000
1r	1.000000e-04	reg	2.500000e+01	batch	750	iters	500	train accuracy:	0.415510	val accuracy:	0.420000
1r	1.000000e-04	reg	2.500000e+01	batch	750	iters	1000	train accuracy:	0.414980	val accuracy:	0.415000
1r	1.000000e-04	reg	2.500000e+01	batch	750	iters	2000	train accuracy:	0.415694	val accuracy:	0.420000
1r	1.000000e-04	reg	5.000000e+01	batch	250	iters	250	train accuracy:	0.414633	val accuracy:	0.421000
1r	1.000000e-04	reg	5.000000e+01	batch	250	iters	500	train accuracy:	0.414367	val accuracy:	0.419000
1r	1.000000e-04	reg	5.000000e+01	batch	250	iters	1000	train accuracy:	0.414347	val accuracy:	0.425000
1r	1.000000e-04	reg	5.000000e+01	batch	250	iters	2000	train accuracy:	0.413592	val accuracy:	0.409000
1r	1.000000e-04	reg	5.000000e+01	batch	500	iters	250	train accuracy:	0.416184	val accuracy:	0.416000

lr 1.000000e-04 reg 5.000000e+02 batch 250 iters 2000 train accuracy: 0.410735 val accuracy: 0.419000
 lr 1.000000e-04 reg 5.000000e+02 batch 500 iters 250 train accuracy: 0.409224 val accuracy: 0.399000
 lr 1.000000e-04 reg 5.000000e+02 batch 500 iters 500 train accuracy: 0.412510 val accuracy: 0.399000
 lr 1.000000e-04 reg 5.000000e+02 batch 500 iters 1000 train accuracy: 0.415245 val accuracy: 0.414000
 lr 1.000000e-04 reg 5.000000e+02 batch 500 iters 2000 train accuracy: 0.407163 val accuracy: 0.399000
 lr 1.000000e-04 reg 5.000000e+02 batch 750 iters 250 train accuracy: 0.413041 val accuracy: 0.410000
 lr 1.000000e-04 reg 5.000000e+02 batch 750 iters 500 train accuracy: 0.408408 val accuracy: 0.401000
 lr 1.000000e-04 reg 5.000000e+02 batch 750 iters 1000 train accuracy: 0.411347 val accuracy: 0.422000
 lr 1.000000e-04 reg 5.000000e+02 batch 750 iters 2000 train accuracy: 0.408959 val accuracy: 0.403000
 lr 1.000000e-04 reg 1.000000e+03 batch 250 iters 250 train accuracy: 0.402918 val accuracy: 0.395000
 lr 1.000000e-04 reg 1.000000e+03 batch 250 iters 500 train accuracy: 0.412592 val accuracy: 0.420000
 lr 1.000000e-04 reg 1.000000e+03 batch 250 iters 1000 train accuracy: 0.411776 val accuracy: 0.394000
 lr 1.000000e-04 reg 1.000000e+03 batch 250 iters 2000 train accuracy: 0.398082 val accuracy: 0.396000
 lr 1.000000e-04 reg 1.000000e+03 batch 500 iters 250 train accuracy: 0.412776 val accuracy: 0.401000
 lr 1.000000e-04 reg 1.000000e+03 batch 500 iters 500 train accuracy: 0.397939 val accuracy: 0.375000
 lr 1.000000e-04 reg 1.000000e+03 batch 500 iters 1000 train accuracy: 0.407939 val accuracy: 0.426000
 lr 1.000000e-04 reg 1.000000e+03 batch 500 iters 2000 train accuracy: 0.407857 val accuracy: 0.401000
 lr 1.000000e-04 reg 1.000000e+03 batch 750 iters 250 train accuracy: 0.407367 val accuracy: 0.408000
 lr 1.000000e-04 reg 1.000000e+03 batch 750 iters 500 train accuracy: 0.411408 val accuracy: 0.423000
 lr 1.000000e-04 reg 1.000000e+03 batch 750 iters 1000 train accuracy: 0.400347 val accuracy: 0.399000
 lr 1.000000e-04 reg 1.000000e+03 batch 750 iters 2000 train accuracy: 0.411122 val accuracy: 0.401000
 lr 1.000000e-04 reg 2.500000e+03 batch 250 iters 250 train accuracy: 0.371878 val accuracy: 0.358000
 lr 1.000000e-04 reg 2.500000e+03 batch 250 iters 500 train accuracy: 0.392571 val accuracy: 0.405000
 lr 1.000000e-04 reg 2.500000e+03 batch 250 iters 1000 train accuracy: 0.396898 val accuracy: 0.393000
 lr 1.000000e-04 reg 2.500000e+03 batch 250 iters 2000 train accuracy: 0.376551 val accuracy: 0.377000
 lr 1.000000e-04 reg 2.500000e+03 batch 500 iters 250 train accuracy: 0.393857 val accuracy: 0.374000
 lr 1.000000e-04 reg 2.500000e+03 batch 500 iters 500 train accuracy: 0.388122 val accuracy: 0.386000
 lr 1.000000e-04 reg 2.500000e+03 batch 500 iters 1000 train accuracy: 0.399714 val accuracy: 0.416000
 lr 1.000000e-04 reg 2.500000e+03 batch 500 iters 2000 train accuracy: 0.385327 val accuracy: 0.378000
 lr 1.000000e-04 reg 2.500000e+03 batch 750 iters 250 train accuracy: 0.402673 val accuracy: 0.397000
 lr 1.000000e-04 reg 2.500000e+03 batch 750 iters 500 train accuracy: 0.408041 val accuracy: 0.396000
 lr 1.000000e-04 reg 2.500000e+03 batch 750 iters 1000 train accuracy: 0.400633 val accuracy: 0.405000
 lr 1.000000e-04 reg 2.500000e+03 batch 750 iters 2000 train accuracy: 0.409449 val accuracy: 0.407000
 lr 1.000000e-04 reg 5.000000e+03 batch 250 iters 250 train accuracy: 0.310388 val accuracy: 0.299000
 lr 1.000000e-04 reg 5.000000e+03 batch 250 iters 500 train accuracy: 0.338163 val accuracy: 0.336000
 lr 1.000000e-04 reg 5.000000e+03 batch 250 iters 1000 train accuracy: 0.329673 val accuracy: 0.316000
 lr 1.000000e-04 reg 5.000000e+03 batch 250 iters 2000 train accuracy: 0.328265 val accuracy: 0.343000
 lr 1.000000e-04 reg 5.000000e+03 batch 500 iters 250 train accuracy: 0.352184 val accuracy: 0.337000
 lr 1.000000e-04 reg 5.000000e+03 batch 500 iters 500 train accuracy: 0.361327 val accuracy: 0.340000
 lr 1.000000e-04 reg 5.000000e+03 batch 500 iters 1000 train accuracy: 0.348408 val accuracy: 0.337000
 lr 1.000000e-04 reg 5.000000e+03 batch 500 iters 2000 train accuracy: 0.354388 val accuracy: 0.357000
 lr 1.000000e-04 reg 5.000000e+03 batch 750 iters 250 train accuracy: 0.392939 val accuracy: 0.394000
 lr 1.000000e-04 reg 5.000000e+03 batch 750 iters 500 train accuracy: 0.364653 val accuracy: 0.348000
 lr 1.000000e-04 reg 5.000000e+03 batch 750 iters 1000 train accuracy: 0.380571 val accuracy: 0.383000
 lr 1.000000e-04 reg 5.000000e+03 batch 750 iters 2000 train accuracy: 0.361510 val accuracy: 0.351000
 lr 1.000000e-04 reg 1.000000e+04 batch 250 iters 250 train accuracy: 0.125612 val accuracy: 0.128000
 lr 1.000000e-04 reg 1.000000e+04 batch 250 iters 500 train accuracy: 0.103143 val accuracy: 0.106000
 lr 1.000000e-04 reg 1.000000e+04 batch 250 iters 1000 train accuracy: 0.094061 val accuracy: 0.093000
 lr 1.000000e-04 reg 1.000000e+04 batch 250 iters 2000 train accuracy: 0.117347 val accuracy: 0.107000
 lr 1.000000e-04 reg 1.000000e+04 batch 500 iters 250 train accuracy: 0.098592 val accuracy: 0.095000
 lr 1.000000e-04 reg 1.000000e+04 batch 500 iters 500 train accuracy: 0.119837 val accuracy: 0.103000
 lr 1.000000e-04 reg 1.000000e+04 batch 500 iters 1000 train accuracy: 0.103306 val accuracy: 0.116000
 lr 1.000000e-04 reg 1.000000e+04 batch 500 iters 2000 train accuracy: 0.075204 val accuracy: 0.069000
 lr 1.000000e-04 reg 1.000000e+04 batch 750 iters 250 train accuracy: 0.100531 val accuracy: 0.087000
 lr 1.000000e-04 reg 1.000000e+04 batch 750 iters 500 train accuracy: 0.105551 val accuracy: 0.101000
 lr 1.000000e-04 reg 1.000000e+04 batch 750 iters 1000 train accuracy: 0.147980 val accuracy: 0.157000
 lr 1.000000e-04 reg 1.000000e+04 batch 750 iters 2000 train accuracy: 0.094367 val accuracy: 0.116000
 lr 1.000000e-04 reg 2.500000e+04 batch 250 iters 250 train accuracy: 0.112000 val accuracy: 0.114000
 lr 1.000000e-04 reg 2.500000e+04 batch 250 iters 500 train accuracy: 0.094959 val accuracy: 0.093000
 lr 1.000000e-04 reg 2.500000e+04 batch 250 iters 1000 train accuracy: 0.100265 val accuracy: 0.087000
 lr 1.000000e-04 reg 2.500000e+04 batch 250 iters 2000 train accuracy: 0.100265 val accuracy: 0.087000
 lr 1.000000e-04 reg 2.500000e+04 batch 500 iters 250 train accuracy: 0.093429 val accuracy: 0.084000
 lr 1.000000e-04 reg 2.500000e+04 batch 500 iters 500 train accuracy: 0.112531 val accuracy: 0.102000
 lr 1.000000e-04 reg 2.500000e+04 batch 500 iters 1000 train accuracy: 0.100265 val accuracy: 0.087000
 lr 1.000000e-04 reg 2.500000e+04 batch 500 iters 2000 train accuracy: 0.100265 val accuracy: 0.087000
 lr 1.000000e-04 reg 2.500000e+04 batch 750 iters 250 train accuracy: 0.100265 val accuracy: 0.089000
 lr 1.000000e-04 reg 2.500000e+04 batch 750 iters 500 train accuracy: 0.087163 val accuracy: 0.088000
 lr 1.000000e-04 reg 2.500000e+04 batch 750 iters 1000 train accuracy: 0.100265 val accuracy: 0.087000

[illegible]

lr 1.000000e-04	reg	2.500000e+06	batch	500	iters	2000	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	2.500000e+06	batch	750	iters	250	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	2.500000e+06	batch	750	iters	500	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	2.500000e+06	batch	750	iters	1000	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	2.500000e+06	batch	750	iters	2000	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	5.000000e+06	batch	250	iters	250	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	5.000000e+06	batch	250	iters	500	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	5.000000e+06	batch	250	iters	1000	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	5.000000e+06	batch	250	iters	2000	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	5.000000e+06	batch	500	iters	250	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	5.000000e+06	batch	500	iters	500	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	5.000000e+06	batch	500	iters	1000	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	5.000000e+06	batch	500	iters	2000	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	5.000000e+06	batch	750	iters	250	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	5.000000e+06	batch	750	iters	500	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	5.000000e+06	batch	750	iters	1000	train accuracy: 0.100265	val accuracy: 0.087000
lr 1.000000e-04	reg	5.000000e+06	batch	750	iters	2000	train accuracy: 0.100265	val accuracy: 0.087000
lr 2.500000e-04	reg	1.000000e+01	batch	250	iters	250	train accuracy: 0.416980	val accuracy: 0.418000
lr 2.500000e-04	reg	1.000000e+01	batch	250	iters	500	train accuracy: 0.427694	val accuracy: 0.431000
lr 2.500000e-04	reg	1.000000e+01	batch	250	iters	1000	train accuracy: 0.427673	val accuracy: 0.434000
lr 2.500000e-04	reg	1.000000e+01	batch	250	iters	2000	train accuracy: 0.431143	val accuracy: 0.436000
lr 2.500000e-04	reg	1.000000e+01	batch	500	iters	250	train accuracy: 0.418061	val accuracy: 0.425000
lr 2.500000e-04	reg	1.000000e+01	batch	500	iters	500	train accuracy: 0.425429	val accuracy: 0.432000
lr 2.500000e-04	reg	1.000000e+01	batch	500	iters	1000	train accuracy: 0.430939	val accuracy: 0.434000
lr 2.500000e-04	reg	1.000000e+01	batch	500	iters	2000	train accuracy: 0.430490	val accuracy: 0.437000
lr 2.500000e-04	reg	1.000000e+01	batch	750	iters	250	train accuracy: 0.420714	val accuracy: 0.421000
lr 2.500000e-04	reg	1.000000e+01	batch	750	iters	500	train accuracy: 0.427551	val accuracy: 0.439000
lr 2.500000e-04	reg	1.000000e+01	batch	750	iters	1000	train accuracy: 0.432408	val accuracy: 0.434000
lr 2.500000e-04	reg	1.000000e+01	batch	750	iters	2000	train accuracy: 0.430796	val accuracy: 0.438000
lr 2.500000e-04	reg	2.500000e+01	batch	250	iters	250	train accuracy: 0.412347	val accuracy: 0.410000
lr 2.500000e-04	reg	2.500000e+01	batch	250	iters	500	train accuracy: 0.413857	val accuracy: 0.411000
lr 2.500000e-04	reg	2.500000e+01	batch	250	iters	1000	train accuracy: 0.416265	val accuracy: 0.425000
lr 2.500000e-04	reg	2.500000e+01	batch	250	iters	2000	train accuracy: 0.412551	val accuracy: 0.416000
lr 2.500000e-04	reg	2.500000e+01	batch	500	iters	250	train accuracy: 0.413653	val accuracy: 0.419000
lr 2.500000e-04	reg	2.500000e+01	batch	500	iters	500	train accuracy: 0.415102	val accuracy: 0.418000
lr 2.500000e-04	reg	2.500000e+01	batch	500	iters	1000	train accuracy: 0.415388	val accuracy: 0.412000
lr 2.500000e-04	reg	2.500000e+01	batch	500	iters	2000	train accuracy: 0.413265	val accuracy: 0.426000
lr 2.500000e-04	reg	2.500000e+01</						

lr 2.500000e-04	reg	2.500000e+02	batch	250	iters	2000	train accuracy: 0.414714	val accuracy: 0.415000
lr 2.500000e-04	reg	2.500000e+02	batch	500	iters	250	train accuracy: 0.407857	val accuracy: 0.396000
lr 2.500000e-04	reg	2.500000e+02	batch	500	iters	500	train accuracy: 0.417306	val accuracy: 0.424000
lr 2.500000e-04	reg	2.500000e+02	batch	500	iters	1000	train accuracy: 0.413408	val accuracy: 0.418000
lr 2.500000e-04	reg	2.500000e+02	batch	500	iters	2000	train accuracy: 0.410224	val accuracy: 0.412000
lr 2.500000e-04	reg	2.500000e+02	batch	750	iters	250	train accuracy: 0.412265	val accuracy: 0.405000
lr 2.500000e-04	reg	2.500000e+02	batch	750	iters	500	train accuracy: 0.414633	val accuracy: 0.413000
lr 2.500000e-04	reg	2.500000e+02	batch	750	iters	1000	train accuracy: 0.414245	val accuracy: 0.430000
lr 2.500000e-04	reg	2.500000e+02	batch	750	iters	2000	train accuracy: 0.411612	val accuracy: 0.416000
lr 2.500000e-04	reg	5.000000e+02	batch	250	iters	250	train accuracy: 0.385816	val accuracy: 0.405000
lr 2.500000e-04	reg	5.000000e+02	batch	250	iters	500	train accuracy: 0.410184	val accuracy: 0.404000
lr 2.500000e-04	reg	5.000000e+02	batch	250	iters	1000	train accuracy: 0.398245	val accuracy: 0.395000
lr 2.500000e-04	reg	5.000000e+02	batch	250	iters	2000	train accuracy: 0.404204	val accuracy: 0.403000
lr 2.500000e-04	reg	5.000000e+02	batch	500	iters	250	train accuracy: 0.402408	val accuracy: 0.408000
lr 2.500000e-04	reg	5.000000e+02	batch	500	iters	500	train accuracy: 0.408041	val accuracy: 0.394000
lr 2.500000e-04	reg	5.000000e+02	batch	500	iters	1000	train accuracy: 0.405061	val accuracy: 0.391000
lr 2.500000e-04	reg	5.000000e+02	batch	500	iters	2000	train accuracy: 0.409531	val accuracy: 0.415000
lr 2.500000e-04	reg	5.000000e+02	batch	750	iters	250	train accuracy: 0.411755	val accuracy: 0.402000
lr 2.500000e-04	reg	5.000000e+02	batch	750	iters	500	train accuracy: 0.417224	val accuracy: 0.415000
lr 2.500000e-04	reg	5.000000e+02	batch	750	iters	1000	train accuracy: 0.418449	val accuracy: 0.414000
lr 2.500000e-04	reg	5.000000e+02	batch	750	iters	2000	train accuracy: 0.407306	val accuracy: 0.392000
lr 2.500000e-04	reg	1.000000e+03	batch	250	iters	250	train accuracy: 0.380286	val accuracy: 0.376000
lr 2.500000e-04	reg	1.000000e+03	batch	250	iters	500	train accuracy: 0.389449	val accuracy: 0.385000
lr 2.500000e-04	reg	1.000000e+03	batch	250	iters	1000	train accuracy: 0.378143	val accuracy: 0.361000
lr 2.500000e-04	reg	1.000000e+03	batch	250	iters	2000	train accuracy: 0.370102	val accuracy: 0.363000
lr 2.500000e-04	reg	1.000000e+03	batch	500	iters	250	train accuracy: 0.385755	val accuracy: 0.396000
lr 2.500000e-04	reg	1.000000e+03	batch	500	iters	500	train accuracy: 0.382061	val accuracy: 0.375000
lr 2.500000e-04	reg	1.000000e+03	batch	500	iters	1000	train accuracy: 0.380184	val accuracy: 0.377000
lr 2.500000e-04	reg	1.000000e+03	batch	500	iters	2000	train accuracy: 0.405000	val accuracy: 0.392000
lr 2.500000e-04	reg	1.000000e+03	batch	750	iters	250	train accuracy: 0.393714	val accuracy: 0.375000
lr 2.500000e-04	reg	1.000000e+03	batch	750	iters	500	train accuracy: 0.399980	val accuracy: 0.402000
lr 2.500000e-04	reg	1.000000e+03	batch	750	iters	1000	train accuracy: 0.407959	val accuracy: 0.405000
lr 2.500000e-04	reg	1.000000e+03	batch	750	iters	2000	train accuracy: 0.402776	val accuracy: 0.406000
lr 2.500000e-04	reg	2.500000e+03	batch	250	iters	250	train accuracy: 0.308776	val accuracy: 0.328000
lr 2.500000e-04	reg	2.500000e+03	batch	250	iters	500	train accuracy: 0.313898	val accuracy: 0.292000
lr 2.500000e-04	reg	2.500000e+03	batch	250	iters	1000	train accuracy: 0.305388	val accuracy: 0.307000
lr 2.500000e-04	reg	2.500000e+03	batch	250	iters	2000	train accuracy: 0.297918	val accuracy: 0.279000
lr 2.500000e-04	reg	2.500000e+03</						

[illegible]

[illegible]

1r 5.000000e-04	reg	1.000000e+02	batch	250	iters	2000	train accuracy: 0.411531	val accuracy: 0.427000
1r 5.000000e-04	reg	1.000000e+02	batch	500	iters	250	train accuracy: 0.417653	val accuracy: 0.424000
1r 5.000000e-04	reg	1.000000e+02	batch	500	iters	500	train accuracy: 0.413367	val accuracy: 0.410000
1r 5.000000e-04	reg	1.000000e+02	batch	500	iters	1000	train accuracy: 0.413388	val accuracy: 0.420000
1r 5.000000e-04	reg	1.000000e+02	batch	500	iters	2000	train accuracy: 0.416878	val accuracy: 0.422000
1r 5.000000e-04	reg	1.000000e+02	batch	750	iters	250	train accuracy: 0.412980	val accuracy: 0.415000
1r 5.000000e-04	reg	1.000000e+02	batch	750	iters	500	train accuracy: 0.413776	val accuracy: 0.422000
1r 5.000000e-04	reg	1.000000e+02	batch	750	iters	1000	train accuracy: 0.411393	val accuracy: 0.400000
1r 5.000000e-04	reg	1.000000e+02	batch	750	iters	2000	train accuracy: 0.411204	val accuracy: 0.404000
1r 5.000000e-04	reg	2.500000e+02	batch	250	iters	250	train accuracy: 0.408347	val accuracy: 0.411000
1r 5.000000e-04	reg	2.500000e+02	batch	250	iters	500	train accuracy: 0.398714	val accuracy: 0.381000
1r 5.000000e-04	reg	2.500000e+02	batch	250	iters	1000	train accuracy: 0.389265	val accuracy: 0.395000
1r 5.000000e-04	reg	2.500000e+02	batch	250	iters	2000	train accuracy: 0.402082	val accuracy: 0.407000
1r 5.000000e-04	reg	2.500000e+02	batch	500	iters	250	train accuracy: 0.417959	val accuracy: 0.425000
1r 5.000000e-04	reg	2.500000e+02	batch	500	iters	500	train accuracy: 0.404796	val accuracy: 0.407000
1r 5.000000e-04	reg	2.500000e+02	batch	500	iters	1000	train accuracy: 0.405796	val accuracy: 0.400000
1r 5.000000e-04	reg	2.500000e+02	batch	500	iters	2000	train accuracy: 0.402082	val accuracy: 0.385000
1r 5.000000e-04	reg	2.500000e+02	batch	750	iters	250	train accuracy: 0.412347	val accuracy: 0.420000
1r 5.000000e-04	reg	2.500000e+02	batch	750	iters	500	train accuracy: 0.403204	val accuracy: 0.410000
1r 5.000000e-04	reg	2.500000e+02	batch	750	iters	1000	train accuracy: 0.406939	val accuracy: 0.412000
1r 5.000000e-04	reg	2.500000e+02	batch	750	iters	2000	train accuracy: 0.409714	val accuracy: 0.409000
1r 5.000000e-04	reg	5.000000e+02	batch	250	iters	250	train accuracy: 0.380878	val accuracy: 0.376000
1r 5.000000e-04	reg	5.000000e+02	batch	250	iters	500	train accuracy: 0.367429	val accuracy: 0.379000
1r 5.000000e-04	reg	5.000000e+02	batch	250	iters	1000	train accuracy: 0.396122	val accuracy: 0.403000
1r 5.000000e-04	reg	5.000000e+02	batch	250	iters	2000	train accuracy: 0.357102	val accuracy: 0.376000
1r 5.000000e-04	reg	5.000000e+02	batch	500	iters	250	train accuracy: 0.396612	val accuracy: 0.390000
1r 5.000000e-04	reg	5.000000e+02	batch	500	iters	500	train accuracy: 0.393755	val accuracy: 0.406000
1r 5.000000e-04	reg	5.000000e+02	batch	500	iters	1000	train accuracy: 0.393531	val accuracy: 0.379000
1r 5.000000e-04	reg	5.000000e+02	batch	500	iters	2000	train accuracy: 0.402245	val accuracy: 0.402000
1r 5.000000e-04	reg	5.000000e+02	batch	750	iters	250	train accuracy: 0.404041	val accuracy: 0.400000
1r 5.000000e-04	reg	5.000000e+02	batch	750	iters	500	train accuracy: 0.393653	val accuracy: 0.382000
1r 5.000000e-04	reg	5.000000e+02	batch	750	iters	1000	train accuracy: 0.411510	val accuracy: 0.395000
1r 5.000000e-04	reg	5.000000e+02	batch	750	iters	2000	train accuracy: 0.405796	val accuracy: 0.407000
1r 5.000000e-04	reg	1.000000e+03	batch	250	iters	250	train accuracy: 0.321163	val accuracy: 0.312000
1r 5.000000e-04	reg	1.000000e+03	batch	250	iters	500	train accuracy: 0.333980	val accuracy: 0.360000
1r 5.000000e-04	reg	1.000000e+03	batch	250	iters	1000	train accuracy: 0.362918	val accuracy: 0.379000
1r 5.000000e-04	reg	1.000000e+03	batch	250	iters	2000	train accuracy: 0.328020	val accuracy: 0.346000
1								

[illegible]

[illegible]

1r	1.000000e-03	reg	5.000000e+01	batch	250	iters	2000	train accuracy: 0.408592	val accuracy: 0.434000
1r	1.000000e-03	reg	5.000000e+01	batch	500	iters	250	train accuracy: 0.413102	val accuracy: 0.421000
1r	1.000000e-03	reg	5.000000e+01	batch	500	iters	500	train accuracy: 0.411388	val accuracy: 0.411000
1r	1.000000e-03	reg	5.000000e+01	batch	500	iters	1000	train accuracy: 0.411592	val accuracy: 0.424000
1r	1.000000e-03	reg	5.000000e+01	batch	500	iters	2000	train accuracy: 0.406041	val accuracy: 0.394000
1r	1.000000e-03	reg	5.000000e+01	batch	750	iters	250	train accuracy: 0.418571	val accuracy: 0.418000
1r	1.000000e-03	reg	5.000000e+01	batch	750	iters	500	train accuracy: 0.415429	val accuracy: 0.413000
1r	1.000000e-03	reg	5.000000e+01	batch	750	iters	1000	train accuracy: 0.414816	val accuracy: 0.407000
1r	1.000000e-03	reg	5.000000e+01	batch	750	iters	2000	train accuracy: 0.415898	val accuracy: 0.415000
1r	1.000000e-03	reg	1.000000e+02	batch	250	iters	250	train accuracy: 0.403143	val accuracy: 0.388000
1r	1.000000e-03	reg	1.000000e+02	batch	250	iters	500	train accuracy: 0.407347	val accuracy: 0.416000
1r	1.000000e-03	reg	1.000000e+02	batch	250	iters	1000	train accuracy: 0.399551	val accuracy: 0.403000
1r	1.000000e-03	reg	1.000000e+02	batch	250	iters	2000	train accuracy: 0.398898	val accuracy: 0.395000
1r	1.000000e-03	reg	1.000000e+02	batch	500	iters	250	train accuracy: 0.412980	val accuracy: 0.396000
1r	1.000000e-03	reg	1.000000e+02	batch	500	iters	500	train accuracy: 0.409469	val accuracy: 0.412000
1r	1.000000e-03	reg	1.000000e+02	batch	500	iters	1000	train accuracy: 0.408735	val accuracy: 0.414000
1r	1.000000e-03	reg	1.000000e+02	batch	500	iters	2000	train accuracy: 0.394224	val accuracy: 0.392000
1r	1.000000e-03	reg	1.000000e+02	batch	750	iters	250	train accuracy: 0.412102	val accuracy: 0.413000
1r	1.000000e-03	reg	1.000000e+02	batch	750	iters	500	train accuracy: 0.407673	val accuracy: 0.408000
1r	1.000000e-03	reg	1.000000e+02	batch	750	iters	1000	train accuracy: 0.410673	val accuracy: 0.428000
1r	1.000000e-03	reg	1.000000e+02	batch	750	iters	2000	train accuracy: 0.407755	val accuracy: 0.413000
1r	1.000000e-03	reg	2.500000e+02	batch	250	iters	250	train accuracy: 0.377490	val accuracy: 0.362000
1r	1.000000e-03	reg	2.500000e+02	batch	250	iters	500	train accuracy: 0.388490	val accuracy: 0.387000
1r	1.000000e-03	reg	2.500000e+02	batch	250	iters	1000	train accuracy: 0.374735	val accuracy: 0.387000
1r	1.000000e-03	reg	2.500000e+02	batch	250	iters	2000	train accuracy: 0.381694	val accuracy: 0.390000
1r	1.000000e-03	reg	2.500000e+02	batch	500	iters	250	train accuracy: 0.387490	val accuracy: 0.379000
1r	1.000000e-03	reg	2.500000e+02	batch	500	iters	500	train accuracy: 0.406449	val accuracy: 0.413000
1r	1.000000e-03	reg	2.500000e+02	batch	500	iters	1000	train accuracy: 0.408408	val accuracy: 0.400000
1r	1.000000e-03	reg	2.500000e+02	batch	500	iters	2000	train accuracy: 0.401347	val accuracy: 0.407000
1r	1.000000e-03	reg	2.500000e+02	batch	750	iters	250	train accuracy: 0.393939	val accuracy: 0.406000
1r	1.000000e-03	reg	2.500000e+02	batch	750	iters	500	train accuracy: 0.415306	val accuracy: 0.422000
1r	1.000000e-03	reg	2.500000e+02	batch	750	iters	1000	train accuracy: 0.403816	val accuracy: 0.394000
1r	1.000000e-03	reg	2.500000e+02	batch	750	iters	2000	train accuracy: 0.382939	val accuracy: 0.386000
1r	1.000000e-03	reg	5.000000e+02	batch	250	iters	250	train accuracy: 0.336918	val accuracy: 0.315000
1r	1.000000e-03	reg	5.000000e+02	batch	250	iters	500	train accuracy: 0.333041	val accuracy: 0.360000
1r	1.000000e-03	reg	5.000000e+02	batch	250	iters	1000	train accuracy: 0.	

[illegible]

[illegible]

lr 2.500000e-03	reg 2.500000e+01	batch 250	iters 2000	train accuracy: 0.407327	val accuracy: 0.412000
lr 2.500000e-03	reg 2.500000e+01	batch 500	iters 250	train accuracy: 0.412041	val accuracy: 0.409000
lr 2.500000e-03	reg 2.500000e+01	batch 500	iters 500	train accuracy: 0.409490	val accuracy: 0.410000
lr 2.500000e-03	reg 2.500000e+01	batch 500	iters 1000	train accuracy: 0.411224	val accuracy: 0.408000
lr 2.500000e-03	reg 2.500000e+01	batch 500	iters 2000	train accuracy: 0.410776	val accuracy: 0.421000
lr 2.500000e-03	reg 2.500000e+01	batch 750	iters 250	train accuracy: 0.409020	val accuracy: 0.404000
lr 2.500000e-03	reg 2.500000e+01	batch 750	iters 500	train accuracy: 0.410449	val accuracy: 0.404000
lr 2.500000e-03	reg 2.500000e+01	batch 750	iters 1000	train accuracy: 0.414429	val accuracy: 0.405000
lr 2.500000e-03	reg 2.500000e+01	batch 750	iters 2000	train accuracy: 0.408204	val accuracy: 0.409000
lr 2.500000e-03	reg 5.000000e+01	batch 250	iters 250	train accuracy: 0.388653	val accuracy: 0.380000
lr 2.500000e-03	reg 5.000000e+01	batch 250	iters 500	train accuracy: 0.394796	val accuracy: 0.407000
lr 2.500000e-03	reg 5.000000e+01	batch 250	iters 1000	train accuracy: 0.397592	val accuracy: 0.388000
lr 2.500000e-03	reg 5.000000e+01	batch 250	iters 2000	train accuracy: 0.391755	val accuracy: 0.389000
lr 2.500000e-03	reg 5.000000e+01	batch 500	iters 250	train accuracy: 0.412898	val accuracy: 0.401000
lr 2.500000e-03	reg 5.000000e+01	batch 500	iters 500	train accuracy: 0.409490	val accuracy: 0.421000
lr 2.500000e-03	reg 5.000000e+01	batch 500	iters 1000	train accuracy: 0.409082	val accuracy: 0.401000
lr 2.500000e-03	reg 5.000000e+01	batch 500	iters 2000	train accuracy: 0.388878	val accuracy: 0.371000
lr 2.500000e-03	reg 5.000000e+01	batch 750	iters 250	train accuracy: 0.403327	val accuracy: 0.416000
lr 2.500000e-03	reg 5.000000e+01	batch 750	iters 500	train accuracy: 0.404429	val accuracy: 0.398000
lr 2.500000e-03	reg 5.000000e+01	batch 750	iters 1000	train accuracy: 0.401816	val accuracy: 0.398000
lr 2.500000e-03	reg 5.000000e+01	batch 750	iters 2000	train accuracy: 0.408286	val accuracy: 0.398000
lr 2.500000e-03	reg 1.000000e+02	batch 250	iters 250	train accuracy: 0.371245	val accuracy: 0.384000
lr 2.500000e-03	reg 1.000000e+02	batch 250	iters 500	train accuracy: 0.369714	val accuracy: 0.384000
lr 2.500000e-03	reg 1.000000e+02	batch 250	iters 1000	train accuracy: 0.372816	val accuracy: 0.375000
lr 2.500000e-03	reg 1.000000e+02	batch 250	iters 2000	train accuracy: 0.373918	val accuracy: 0.379000
lr 2.500000e-03	reg 1.000000e+02	batch 500	iters 250	train accuracy: 0.383327	val accuracy: 0.384000
lr 2.500000e-03	reg 1.000000e+02	batch 500	iters 500	train accuracy: 0.395796	val accuracy: 0.416000
lr 2.500000e-03	reg 1.000000e+02	batch 500	iters 1000	train accuracy: 0.397796	val accuracy: 0.398000
lr 2.500000e-03	reg 1.000000e+02	batch 500	iters 2000	train accuracy: 0.375551	val accuracy: 0.354000
lr 2.500000e-03	reg 1.000000e+02	batch 750	iters 250	train accuracy: 0.386694	val accuracy: 0.375000
lr 2.500000e-03	reg 1.000000e+02	batch 750	iters 500	train accuracy: 0.389898	val accuracy: 0.372000
lr 2.500000e-03	reg 1.000000e+02	batch 750	iters 1000	train accuracy: 0.401959	val accuracy: 0.423000
lr 2.500000e-03	reg 1.000000e+02	batch 750	iters 2000	train accuracy: 0.395265	val accuracy: 0.387000
lr 2.500000e-03	reg 2.500000e+02	batch 250	iters 250	train accuracy: 0.294633	val accuracy: 0.274000
lr 2.500000e-03	reg 2.500000e+02	batch 250	iters 500	train accuracy: 0.301469	val accuracy: 0.304000
lr 2.500000e-03	reg 2.500000e+02	batch 250	iters 1000	train accuracy: 0.294388	val accuracy: 0.287000
lr 2.500000e-03	reg 2.500000e+02	batch 250	iters 2000	train accuracy: 0.325082	val accuracy: 0.302000
lr 2.500000e-03	reg 2.500000e+02	batch 500	iters 250	train accuracy: 0.332612	val accuracy: 0.314000
lr 2.500000e-03	reg 2.500000e+02	batch 500	iters 500	train accuracy: 0.347592	val accuracy: 0.351000
lr 2.500000e-03	reg 2.500000e+02	batch 500	iters 1000	train accuracy: 0.353776	val accuracy: 0.356000
lr 2.500000e-03	reg 2.500000e+02	batch 500	iters 2000	train accuracy: 0.322204	val accuracy: 0.333000
lr 2.500000e-03	reg 2.500000e+02	batch 750	iters 250	train accuracy: 0.359449	val accuracy: 0.362000

[illegible]

[illegible]

1r 5.000000e-03	reg	1.000000e+01	batch	250	iters	2000	train accuracy: 0.430592	val accuracy: 0.433000
1r 5.000000e-03	reg	1.000000e+01	batch	500	iters	250	train accuracy: 0.428327	val accuracy: 0.417000
1r 5.000000e-03	reg	1.000000e+01	batch	500	iters	500	train accuracy: 0.428327	val accuracy: 0.420000
1r 5.000000e-03	reg	1.000000e+01	batch	500	iters	1000	train accuracy: 0.426122	val accuracy: 0.426000
1r 5.000000e-03	reg	1.000000e+01	batch	500	iters	2000	train accuracy: 0.427510	val accuracy: 0.433000
1r 5.000000e-03	reg	1.000000e+01	batch	750	iters	250	train accuracy: 0.426286	val accuracy: 0.421000
1r 5.000000e-03	reg	1.000000e+01	batch	750	iters	500	train accuracy: 0.427327	val accuracy: 0.433000
1r 5.000000e-03	reg	1.000000e+01	batch	750	iters	1000	train accuracy: 0.430878	val accuracy: 0.438000
1r 5.000000e-03	reg	1.000000e+01	batch	750	iters	2000	train accuracy: 0.429694	val accuracy: 0.422000
1r 5.000000e-03	reg	2.500000e+01	batch	250	iters	250	train accuracy: 0.389429	val accuracy: 0.377000
1r 5.000000e-03	reg	2.500000e+01	batch	250	iters	500	train accuracy: 0.399429	val accuracy: 0.386000
1r 5.000000e-03	reg	2.500000e+01	batch	250	iters	1000	train accuracy: 0.393245	val accuracy: 0.419000
1r 5.000000e-03	reg	2.500000e+01	batch	250	iters	2000	train accuracy: 0.404245	val accuracy: 0.405000
1r 5.000000e-03	reg	2.500000e+01	batch	500	iters	250	train accuracy: 0.404082	val accuracy: 0.417000
1r 5.000000e-03	reg	2.500000e+01	batch	500	iters	500	train accuracy: 0.395082	val accuracy: 0.385000
1r 5.000000e-03	reg	2.500000e+01	batch	500	iters	1000	train accuracy: 0.407490	val accuracy: 0.411000
1r 5.000000e-03	reg	2.500000e+01	batch	500	iters	2000	train accuracy: 0.407469	val accuracy: 0.399000
1r 5.000000e-03	reg	2.500000e+01	batch	750	iters	250	train accuracy: 0.410714	val accuracy: 0.412000
1r 5.000000e-03	reg	2.500000e+01	batch	750	iters	500	train accuracy: 0.397816	val accuracy: 0.390000
1r 5.000000e-03	reg	2.500000e+01	batch	750	iters	1000	train accuracy: 0.411204	val accuracy: 0.407000
1r 5.000000e-03	reg	2.500000e+01	batch	750	iters	2000	train accuracy: 0.402306	val accuracy: 0.404000
1r 5.000000e-03	reg	5.000000e+01	batch	250	iters	250	train accuracy: 0.394408	val accuracy: 0.385000
1r 5.000000e-03	reg	5.000000e+01	batch	250	iters	500	train accuracy: 0.356163	val accuracy: 0.358000
1r 5.000000e-03	reg	5.000000e+01	batch	250	iters	1000	train accuracy: 0.368245	val accuracy: 0.356000
1r 5.000000e-03	reg	5.000000e+01	batch	250	iters	2000	train accuracy: 0.368633	val accuracy: 0.378000
1r 5.000000e-03	reg	5.000000e+01	batch	500	iters	250	train accuracy: 0.386388	val accuracy: 0.368000
1r 5.000000e-03	reg	5.000000e+01	batch	500	iters	500	train accuracy: 0.383449	val accuracy: 0.381000
1r 5.000000e-03	reg	5.000000e+01	batch	500	iters	1000	train accuracy: 0.393510	val accuracy: 0.389000
1r 5.000000e-03	reg	5.000000e+01	batch	500	iters	2000	train accuracy: 0.382388	val accuracy: 0.368000
1r 5.000000e-03	reg	5.000000e+01	batch	750	iters	250	train accuracy: 0.398510	val accuracy: 0.390000
1r 5.000000e-03	reg	5.000000e+01	batch	750	iters	500	train accuracy: 0.407286	val accuracy: 0.412000
1r 5.000000e-03	reg	5.000000e+01	batch	750	iters	1000	train accuracy: 0.399653	val accuracy: 0.405000
1r 5.000000e-03	reg	5.000000e+01	batch	750	iters	2000	train accuracy: 0.399408	val accuracy: 0.398000
1r 5.000000e-03	reg	1.000000e+02	batch	250	iters	250	train accuracy: 0.340878	val accuracy: 0.349000
1r 5.000000e-03	reg	1.000000e+02	batch	250	iters	500	train accuracy: 0.338735	val accuracy: 0.319000
1r 5.000000e-03	reg	1.000000e+02	batch	250	iters	1000	train accuracy: 0.339673	val accuracy: 0.342000
1r 5.000000e-03	reg	1.000000e+02	batch	250	iters	2000	train accuracy: 0.335592	val accuracy: 0.349000
1								

[illegible]

[illegible]

```

lr 5.000000e-03 reg 5.000000e+06 batch 250 iters 2000 train accuracy: 0.100265 val accuracy: 0.087000
lr 5.000000e-03 reg 5.000000e+06 batch 500 iters 250 train accuracy: 0.100265 val accuracy: 0.087000
lr 5.000000e-03 reg 5.000000e+06 batch 500 iters 500 train accuracy: 0.100265 val accuracy: 0.087000
lr 5.000000e-03 reg 5.000000e+06 batch 500 iters 1000 train accuracy: 0.100265 val accuracy: 0.087000
lr 5.000000e-03 reg 5.000000e+06 batch 500 iters 2000 train accuracy: 0.100265 val accuracy: 0.087000
lr 5.000000e-03 reg 5.000000e+06 batch 750 iters 250 train accuracy: 0.100265 val accuracy: 0.087000
lr 5.000000e-03 reg 5.000000e+06 batch 750 iters 500 train accuracy: 0.100265 val accuracy: 0.087000
lr 5.000000e-03 reg 5.000000e+06 batch 750 iters 1000 train accuracy: 0.100265 val accuracy: 0.087000
lr 5.000000e-03 reg 5.000000e+06 batch 750 iters 2000 train accuracy: 0.100265 val accuracy: 0.087000
best validation accuracy achieved during cross-validation: 0.445000

```

```

In [17]: # Evaluate your trained SVM on the test set
y_test_pred = best_svm.predict(X_test_feats)
test_accuracy = np.mean(y_test == y_test_pred)
print(test_accuracy)

```

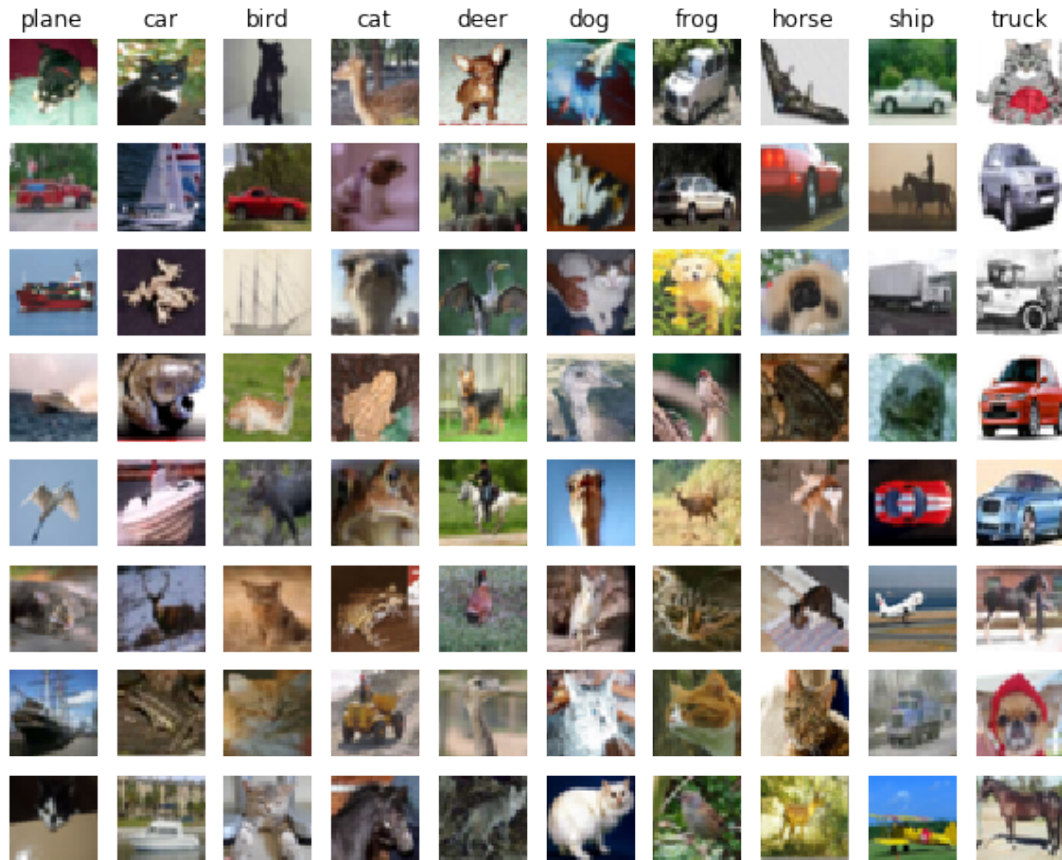
0.439

```

In [18]: # An important way to gain intuition about how an algorithm works is to
# visualize the mistakes that it makes. In this visualization, we show examples
# of images that are misclassified by our current system. The first column
# shows images that our system labeled as "plane" but whose true label is
# something other than "plane".

examples_per_class = 8
classes = ['plane', 'car', 'bird', 'cat', 'deer', 'dog', 'frog', 'horse', 'ship',
'truck']
for cls, cls_name in enumerate(classes):
    idxs = np.where((y_test != cls) & (y_test_pred == cls))[0]
    idxs = np.random.choice(idxs, examples_per_class, replace=False)
    for i, idx in enumerate(idxs):
        plt.subplot(examples_per_class, len(classes), i * len(classes) + cls + 1)
        plt.imshow(X_test[idx].astype('uint8'))
        plt.axis('off')
        if i == 0:
            plt.title(cls_name)
plt.show()

```



1.3.1 Inline question 1:

Describe the misclassification results that you see. Do they make sense?

Some misclassification results are hard to interpret, but there are a number which can be reasonably explained -

1. The images in the “plane” column are all characterized by large proportions of the image being either blue or grey-blue, which suggests that one of the main characteristics of planes that the network learned was that they are often found in the sky.
2. Many of the images in the “deer” column are other animals with protruding appendages (a dog with big ears; a few birds with outstretched wings), suggesting that the neural network learned to classify deer by looking for antlers.
3. Half of the images in the “truck” column are cars, suggesting that the neural network learned something about how to classify four-wheeled passenger vehicles, but perhaps not quite enough to distinguish between sedans and pickup trucks.
4. Interestingly, half of the images in the “dog” column are cats, while only one of the images in the “cat” column is a dog, suggesting that the network learned that some characteristic of a dog is also not characteristic of a cat, but not necessarily any characteristic of either which would suggest the opposite.

1.4 Neural Network on image features

Earlier in this assignment we saw that training a two-layer neural network on raw pixels achieved better classification performance than linear classifiers on raw pixels. In this notebook we have seen that linear classifiers on image features outperform linear classifiers on raw pixels.

For completeness, we should also try training a neural network on image features. This approach should outperform all previous approaches: you should easily be able to achieve over 55% classification accuracy on the test set; our best model achieves about 60% classification accuracy.

```
In [19]: # Preprocessing: Remove the bias dimension
# Make sure to run this cell only ONCE
print(X_train_feats.shape)
X_train_feats = X_train_feats[:, :-1]
X_val_feats = X_val_feats[:, :-1]
X_test_feats = X_test_feats[:, :-1]

print(X_train_feats.shape)

(49000, 155)
(49000, 154)

In [56]: from cs682.classifiers.neural_net import TwoLayerNet

input_size = X_train_feats.shape[1]
hidden_dim = 500
num_classes = 10

#####
# TODO: Train a two-layer neural network on image features. You may want to #
# cross-validate various parameters as in previous sections. Store your best #
# model in the best_net variable. #
#####

def select_hyper_parameters():
    results = {}
    best_acc = -1
    best_net = None

    learning_rates = [5e-2, 1e-1, 2.5e-1, 5e-1, 1.0]
    regularization_strengths = [1e-4, 2.5e-4, 5e-4, 1e-3, 2.5e-3]
    hidden_layer_sizes = [500, 750, 1000]
    batch_sizes = [50, 100]
    num_epochs = 4

    for rate in learning_rates:
        for strength in regularization_strengths:
            for units in hidden_layer_sizes:
                for batch_size in batch_sizes:

                    num_iters = num_epochs * int(X_train_feats.shape[0] / batch_size)

                    net = TwoLayerNet(input_size, units, num_classes)

                    stats = net.train(X_train_feats, y_train, X_val_feats, y_val,
                                     learning_rate=rate,
                                     reg=strength,
                                     batch_size=batch_size,
                                     num_iters=num_iters)

                    val_acc = (net.predict(X_val_feats) == y_val).mean()
                    print('rate: %f, strength: %.7f, hidden units: %d, batch size: %d,
acc: %f' % (rate,
              strength,
              units,
              batch_size,
              val_acc))

                    if val_acc > best_acc:
                        best_acc = val_acc
                        best_net = net

                    results[(rate, strength, units, batch_size, num_iters)] = (val_acc,
stats)

    return best_net, best_acc, results

# rate: 0.250000, strength: 0.0001000, hidden units: 1000, batch size: 100, acc:
```

```

0.607000
best_learning_rate = 2.5e-1
best_reg_strength = 1e-4
best_num_hidden_units = 1000
best_batch_size = 100
num_epochs = 4
perform_cross_validation = False

if perform_cross_validation is True:
    best_net, best_acc, results = select_hyper_parameters()
else:
    best_net = TwoLayerNet(input_size, best_num_hidden_units, num_classes)

    num_iters = num_epochs * int(X_train.shape[0] / best_batch_size)

    stats = net.train(X_train, y_train, X_val, y_val,
                      learning_rate=best_learning_rate,
                      reg=best_reg_strength,
                      batch_size=best_batch_size,
                      num_iters=num_iters)

    best_acc = (best_net.predict(X_val) == y_val).mean()

    print('rate: %f, strength: %.7f, hidden units: %d, batch size: %d, acc: %f' %
          (best_learning_rate,
           best_reg_strength,
           best_num_hidden_units,
           best_batch_size,
           best_acc))

#####
#                               END OF YOUR CODE                               #
#####

rate: 0.050000, strength: 0.0001000, hidden units: 500, batch size: 50, acc: 0.544000
rate: 0.050000, strength: 0.0001000, hidden units: 500, batch size: 100, acc: 0.514000
rate: 0.050000, strength: 0.0001000, hidden units: 750, batch size: 50, acc: 0.549000
rate: 0.050000, strength: 0.0001000, hidden units: 750, batch size: 100, acc: 0.525000
rate: 0.050000, strength: 0.0001000, hidden units: 1000, batch size: 50, acc: 0.526000
rate: 0.050000, strength: 0.0001000, hidden units: 1000, batch size: 100, acc: 0.522000
rate: 0.050000, strength: 0.0002500, hidden units: 500, batch size: 50, acc: 0.539000
rate: 0.050000, strength: 0.0002500, hidden units: 500, batch size: 100, acc: 0.514000
rate: 0.050000, strength: 0.0002500, hidden units: 750, batch size: 50, acc: 0.537000
rate: 0.050000, strength: 0.0002500, hidden units: 750, batch size: 100, acc: 0.511000
rate: 0.050000, strength: 0.0002500, hidden units: 1000, batch size: 50, acc: 0.540000
rate: 0.050000, strength: 0.0002500, hidden units: 1000, batch size: 100, acc: 0.517000
rate: 0.050000, strength: 0.0005000, hidden units: 500, batch size: 50, acc: 0.541000
rate: 0.050000, strength: 0.0005000, hidden units: 500, batch size: 100, acc: 0.513000
rate: 0.050000, strength: 0.0005000, hidden units: 750, batch size: 50, acc: 0.553000
rate: 0.050000, strength: 0.0005000, hidden units: 750, batch size: 100, acc: 0.508000
rate: 0.050000, strength: 0.0005000, hidden units: 1000, batch size: 50, acc: 0.536000
rate: 0.050000, strength: 0.0005000, hidden units: 1000, batch size: 100, acc: 0.520000
rate: 0.050000, strength: 0.0010000, hidden units: 500, batch size: 50, acc: 0.540000
rate: 0.050000, strength: 0.0010000, hidden units: 500, batch size: 100, acc: 0.503000
rate: 0.050000, strength: 0.0010000, hidden units: 750, batch size: 50, acc: 0.531000
rate: 0.050000, strength: 0.0010000, hidden units: 750, batch size: 100, acc: 0.524000
rate: 0.050000, strength: 0.0010000, hidden units: 1000, batch size: 50, acc: 0.528000
rate: 0.050000, strength: 0.0010000, hidden units: 1000, batch size: 100, acc: 0.524000
rate: 0.050000, strength: 0.0025000, hidden units: 500, batch size: 50, acc: 0.525000
rate: 0.050000, strength: 0.0025000, hidden units: 500, batch size: 100, acc: 0.518000
rate: 0.050000, strength: 0.0025000, hidden units: 750, batch size: 50, acc: 0.533000
rate: 0.050000, strength: 0.0025000, hidden units: 750, batch size: 100, acc: 0.513000
rate: 0.050000, strength: 0.0025000, hidden units: 1000, batch size: 50, acc: 0.524000
rate: 0.050000, strength: 0.0025000, hidden units: 1000, batch size: 100, acc: 0.511000
rate: 0.100000, strength: 0.0001000, hidden units: 500, batch size: 50, acc: 0.572000
rate: 0.100000, strength: 0.0001000, hidden units: 500, batch size: 100, acc: 0.545000
rate: 0.100000, strength: 0.0001000, hidden units: 750, batch size: 50, acc: 0.561000
rate: 0.100000, strength: 0.0001000, hidden units: 750, batch size: 100, acc: 0.543000
rate: 0.100000, strength: 0.0001000, hidden units: 1000, batch size: 50, acc: 0.588000
rate: 0.100000, strength: 0.0001000, hidden units: 1000, batch size: 100, acc: 0.527000
rate: 0.100000, strength: 0.0002500, hidden units: 500, batch size: 50, acc: 0.577000
rate: 0.100000, strength: 0.0002500, hidden units: 500, batch size: 100, acc: 0.534000
rate: 0.100000, strength: 0.0002500, hidden units: 750, batch size: 50, acc: 0.572000

```

[illegible]

```

rate: 0.500000, strength: 0.0005000, hidden units: 1000, batch size: 100, acc: 0.567000
rate: 0.500000, strength: 0.0010000, hidden units: 500, batch size: 50, acc: 0.556000
rate: 0.500000, strength: 0.0010000, hidden units: 500, batch size: 100, acc: 0.563000
rate: 0.500000, strength: 0.0010000, hidden units: 750, batch size: 50, acc: 0.551000
rate: 0.500000, strength: 0.0010000, hidden units: 750, batch size: 100, acc: 0.580000
rate: 0.500000, strength: 0.0010000, hidden units: 1000, batch size: 50, acc: 0.556000
rate: 0.500000, strength: 0.0010000, hidden units: 1000, batch size: 100, acc: 0.586000
rate: 0.500000, strength: 0.0025000, hidden units: 500, batch size: 50, acc: 0.536000
rate: 0.500000, strength: 0.0025000, hidden units: 500, batch size: 100, acc: 0.549000
rate: 0.500000, strength: 0.0025000, hidden units: 750, batch size: 50, acc: 0.511000
rate: 0.500000, strength: 0.0025000, hidden units: 750, batch size: 100, acc: 0.545000
rate: 0.500000, strength: 0.0025000, hidden units: 1000, batch size: 50, acc: 0.515000
rate: 0.500000, strength: 0.0025000, hidden units: 1000, batch size: 100, acc: 0.555000
rate: 1.000000, strength: 0.0001000, hidden units: 500, batch size: 50, acc: 0.513000
rate: 1.000000, strength: 0.0001000, hidden units: 500, batch size: 100, acc: 0.543000
rate: 1.000000, strength: 0.0001000, hidden units: 750, batch size: 50, acc: 0.511000
rate: 1.000000, strength: 0.0001000, hidden units: 750, batch size: 100, acc: 0.556000
rate: 1.000000, strength: 0.0001000, hidden units: 1000, batch size: 50, acc: 0.494000
rate: 1.000000, strength: 0.0001000, hidden units: 1000, batch size: 100, acc: 0.563000
rate: 1.000000, strength: 0.0002500, hidden units: 500, batch size: 50, acc: 0.519000
rate: 1.000000, strength: 0.0002500, hidden units: 500, batch size: 100, acc: 0.546000
rate: 1.000000, strength: 0.0002500, hidden units: 750, batch size: 50, acc: 0.518000
rate: 1.000000, strength: 0.0002500, hidden units: 750, batch size: 100, acc: 0.553000
rate: 1.000000, strength: 0.0002500, hidden units: 1000, batch size: 50, acc: 0.482000
rate: 1.000000, strength: 0.0002500, hidden units: 1000, batch size: 100, acc: 0.554000
rate: 1.000000, strength: 0.0005000, hidden units: 500, batch size: 50, acc: 0.503000
rate: 1.000000, strength: 0.0005000, hidden units: 500, batch size: 100, acc: 0.538000
rate: 1.000000, strength: 0.0005000, hidden units: 750, batch size: 50, acc: 0.503000
rate: 1.000000, strength: 0.0005000, hidden units: 750, batch size: 100, acc: 0.550000
rate: 1.000000, strength: 0.0005000, hidden units: 1000, batch size: 50, acc: 0.508000
rate: 1.000000, strength: 0.0005000, hidden units: 1000, batch size: 100, acc: 0.538000
rate: 1.000000, strength: 0.0010000, hidden units: 500, batch size: 50, acc: 0.489000
rate: 1.000000, strength: 0.0010000, hidden units: 500, batch size: 100, acc: 0.544000
rate: 1.000000, strength: 0.0010000, hidden units: 750, batch size: 50, acc: 0.473000
rate: 1.000000, strength: 0.0010000, hidden units: 750, batch size: 100, acc: 0.526000
rate: 1.000000, strength: 0.0010000, hidden units: 1000, batch size: 50, acc: 0.526000
rate: 1.000000, strength: 0.0010000, hidden units: 1000, batch size: 100, acc: 0.536000
rate: 1.000000, strength: 0.0025000, hidden units: 500, batch size: 50, acc: 0.494000
rate: 1.000000, strength: 0.0025000, hidden units: 500, batch size: 100, acc: 0.521000
rate: 1.000000, strength: 0.0025000, hidden units: 750, batch size: 50, acc: 0.487000
rate: 1.000000, strength: 0.0025000, hidden units: 750, batch size: 100, acc: 0.533000
rate: 1.000000, strength: 0.0025000, hidden units: 1000, batch size: 50, acc: 0.488000
rate: 1.000000, strength: 0.0025000, hidden units: 1000, batch size: 100, acc: 0.544000

```

```

In [57]: # Run your best neural net classifier on the test set. You should be able
         # to get more than 55% accuracy.

```

```

test_acc = (best_net.predict(X_test_feats) == y_test).mean()
print(test_acc)

```

```
0.575
```

```
In [ ]:
```