CS 445/545

Machine Learning

Winter 2017

Homework 1: Perceptrons

Due Tuesday, January 24, 2017, 2pm

For this homework you will write code to implement perceptrons and the perceptron learning algorithm.

If you haven't already, please sign up for the class mailing list:

https://mailhost.cecs.pdx.edu/mailman/listinfo/ml2017

Perceptrons

You will use a perceptron with 785 inputs (including bias input) and 10 outputs to learn to classify the handwritten digits in the MNIST dataset. See the class slides for details of the perceptron architecture and perceptron learning algorithm.

Preprocessing: Scale each data value to be between 0 and 1 (i.e., divide each value by 255, which is the maximum value in the original data). This will help keep the weights from getting too large.

Training: Train perceptrons with three different learning rates: $\eta = 0.001$, 0.01, and 0.1.

For each learning rate:

1. Choose small random initial weights, $w_i \in [-.05, .05]$. Compute the accuracy on the training and test sets for this initial set of weights, to include in your plot. (Call this "epoch 0".)

Recall that the bias unit is always set to 1, and the bias weight is treated like any other weight.

2. Repeat: cycle through the training data, changing the weights (according to the perceptron learning rule) after processing each training example $\boldsymbol{x}^k$, as follows:

- Compute $\boldsymbol{w} \cdot \boldsymbol{x}^k$ at each output unit.
- The unit with the highest value of $\boldsymbol{w} \cdot \boldsymbol{x}^k$ is the prediction for this training example.
- If this is the correct prediction, don't change the weights and go on to the next

training example.

- Otherwise, update all weights in the perceptron:

$$w_i \leftarrow w_i + \eta(t^k - y^k)x_i^k,$$

where

$$t^k = \begin{cases} 1 \text{ if the ouput unit is the correct one for this training example} \\ 0 \text{ otherwise} \end{cases}$$

and

$$y^k = \begin{cases} 1 \text{ if } \boldsymbol{w} \cdot \boldsymbol{x}^k > 0 \\ 0 \text{ otherwise} \end{cases}$$

Thus, $(t^k - y^k)$ can be 1, 0, or −1.

(Note that this means that for some output units $(t^k - y^k)$ could be zero, and thus the weights to that output unit would not be updated, even if the prediction was incorrect. That's okay!)

Keep repeating until the accuracy on the training data has essentially stopped improving (i.e., the difference between training accuracy from one epoch to the next is less than some small number, like .01,) or you have run for 70 epochs (iterations through the training set), whichever comes first.

After each epoch (one cycle through training data), compute accuracy on training and test set (for plot), without changing weights.

If your perceptron training does not seem to be converging (e.g., it is oscillating between two fairly different accuracy values) try starting over with a different set of random initial weights. But note that for some values of the learning rate, you will likely observe cycling on any initial set of weights.

Report: Your report should be a one paragraph description of the experiment, including, ***for each learning rate***:

- Plot of accuracy (fraction of correct classifications) on the training and test set at each epoch (including epoch 0), along with comments as to whether you are seeing either oscillations or overfitting.

- Confusion matrix on the test set, after training has been completed.

This homework's report is very short—just a paragraph, plots and confusion matrix.

Here is what you need to turn in:

- Your report, with all the information requested above.
- Your well-commented code.

How to turn it in (read carefully!):

- Send these items in electronic format to mm@pdx.edu by 2pm on the due date. No hard copy please!

- The report should be in pdf format and the code should be in plain-text format. Make sure your name is on both the report and the code.

- Put "MACHINE LEARNING HW 1" in the subject line.

If there are any questions on this assignment, don't hesitate to ask me or e-mail the class mailing list.

Policy on late homework: If you are having trouble completing the assignment on time for any reason, please see me before the due date to find out if you can get an extension. Any homework turned in late without an extension from me will have 5% of the grade subtracted for each day the assignment is late.