CMSC 478 – Introduction to Machine Learning
Spring 2022
Homework #1

Due Wednesday, February 16ᵗʰ, at 11:30 via Slack to the TA

Question 1: Use Scikit

Your first task is to install Scikit. There are a few ways to do that. You can either install Python (https://www.python.org/downloads/) and then install Scikit (https://scikit-learn.org/stable/install.html), or you can install Anaconda (https://docs.anaconda.com/anaconda/install/index.html) and get both Python and Scikit at the same time. For some of you this will be easy, and some of you will run into problems due to random system configuration issues beyond your control. Do the installation early, and ask questions in slack because your classmates will run into and solve some of the same issues.

When you're up and running, read the documentation on the Scikit implementation of the Perceptron algorithm (https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.Perceptron.html).

On the bottom of that page there is an example of a few lines of code that downloads the digits dataset we saw in class and trains a Perceptron. Get and run that code.

Do the following:
1. Count the number of instances of each class that are misclassified by the trained Perceptron. Submit the code that you wrote to get that count and the final counts.
2. Run the Perceptron a second time but set the fit_intercept keyword to False and show the resulting accuracy. What happened to accuracy and why?
3. Print the weights associated with class/digit 0 as an 8x8 array using this line of code and submit the output:
    a. [clf.coef_[0][i:i+8] for i in range(0, 64, 8)]
4. Recall that the inputs consist of 8x8 grayscale digits with a pixel value of 0 being black and more positive values being closer to white. Also recall that negative weights mean that the feature/pixel having a positive value makes the input less likely to be the target class and positive weights mean that the feature/pixel having a positive value makes the input more likely to be the target class. Given that, explain whether it seems like the

weights for class 0 are reasonable and why. What do those weights say about what kinds of images are likely to be class 0?

Question 2: Loss functions and gradient descent

Polynomial regression is like linear regression, but the terms in the equation used to fit the data can be nonlinear functions of the input variables. Suppose you have a dataset where each observation is a single value, and you want to fit the dataset with the following functional form:

$$f(x) = w_2 x^2 + w_1 x + w_0$$

Note that values of x are given in the dataset, and your job is to learn values for the weights that minimize the loss on the dataset. The loss will be squared error, so the objective is to minimize the sum over the entire dataset of the difference between the predicted value and the true value:

$$J(w) = \sum_{i=1}^{n} (y_i - f(x_i))^2$$

Here is a dataset:

x	y
1	-1
2	6
3	17

Do the following:
1. Write down an expression for the objective function for this dataset by plugging in x and y values into the expression above. Note that the weights will be left as variables.
2. Take the partial derivative of that function with respect to each of the weights, yielding three equations. Simplify each equation as much as possible by combining like terms. Feel free to use an online tool to do the simplification, but let me know what tool you used in your answer. Show the resulting equations.
3. You now have three *linear* equations with three unknowns (the weights). Set each equation equal to 0 and solve for the values of the weights. Feel free to use an online

tool to solve the system of equations, but let me know what tool you used in your answer. Show the resulting weights.

4. Suppose your initial guess for each of the weights is 0, and the learning rate is 0.001. Show the weights after each of two steps of gradient descent. You can use the partial derivatives obtained in step 2 above here. Show your work.

5. Explain whether it looks like gradient descent is going in the right or the wrong direction, and why. Say something about each of the weights in your answer.