CS 4033/5033 - Machine Learning: Homework 4

Spring 2022

Due: Wednesday, March 30, 2022 (11:59pm CDT)

Exercise 1 – Nearest Neighbors (20 points). We are given the following training examples in 2D for binary classification:

$$((-3, 5), +), \ \ ((-4, -2), +), \ \ ((2, 1), -), \ \ ((4, 3), -)$$

Assume that we want to classify the points $(-2, 0)$ and $(-1, 5)$ using

(a) **(10 pts)** a 1-nearest neighbor rule, and

(b) **(10 pts)** a 3-nearest neighbor rule.

Use the Euclidean distance (L_2-norm) for calculating the distance between the points, in order to determine the k nearest neighbors to the query points given. What is the classification of these two points in each case?

Exercise 2 – Perceptron (30 points). We are given the following training examples in 2D (same as above):

$$((-3, 5), +), \ \ ((-4, -2), +), \ \ ((2, 1), -), \ \ ((4, 3), -)$$

Use +1 to map positive (+) examples and -1 to map negative (-) examples.

We want to apply the learning algorithm for training a perceptron using the above data with starting weights $w_0 = w_1 = w_2 = 0$ and learning rate $\eta = 0.1$. In each iteration process the training examples in the order given above. Complete at most 3 iterations over the above training examples.

(a) **(24 pts)** What are the weights at the end of each iteration?

(b) **(6 pts)** Are these weights final?

Exercise 3 – Nearest Neighbors (20 points). We are given the following training examples:

$$(1.2, 3.2), \ \ (2.8, 8.5), \ \ (2, 4.7), \ \ (0.9, 2.9), \ \ (5.1, 11)$$

We want to apply a 3-nearest neighbor rule in order to perform regression.

(a) **(10 pts)** Predict the label (real value) at each of the following two points: $x_1 = 1.5$ and $x_2 = 4.5$.

(b) **(10 pts)** Instead of weighing the contribution of each of the 3 nearest neighbors equally, this time we want to perform distance-weighted nearest neighbor regression. What values do we predict now for $x_1 = 1.5$ and $x_2 = 4.5$?

Exercise 4 – Gradient Descent (30 points). We are given the following training examples (same as above):

$$(1.2, 3.2),\quad (2.8, 8.5),\quad (2, 4.7),\quad (0.9, 2.9),\quad (5.1, 11)$$

Suppose the weights are $w_0 = w_1 = 1$ initially. We want to minimize the cumulative loss $\mathcal{L}_S(h, c)$ that corresponds to the half of the residual sum of squares; i.e., we have that

$$\mathcal{L}_S(h, c) = \frac{1}{2} RSS_S(h, c) = \frac{1}{2}\sum_{i=1}^{m}(y_i - h(x_i))^2$$

and we will be using full gradient descent as discussed in our slides (slide 26 from Module 7).

(a) **(24 pts)** Using $\eta = 0.01$, perform three iterations of full gradient descent, listing w_0 and w_1 at each iteration.

(b) **(6 pts)** What is the value that we predict at the following points $x_1 = 1.5$ and $x_2 = 4.5$?