University of Illinois Spring 2018

CS 446: Machine Learning
Homework 1

Due on Tuesday, January 23, 2018, 11:59 a.m. Central Time

1. **[4 points]** Intro to Machine Learning

   Consider the task of classifying an image as one of a set of objects. Suppose we use a convolutional neural network to do so (you will learn what this is later in the semester).

   (a) For this setup, what is the data (often referred to as $x^{(i)}$)?

   Your answer: Each datum $x^{(i)}$ is an input image.

   (b) For this setup, what is the label (often referred to as $y^{(i)}$)?

   Your answer: Each label $y^{(i)}$ is a member of an enumerated set (potentially an integer) representing the classification of its corresponding example.

   (c) For this setup, what is the model?

   Your answer: The model describes the various layers in the network, what operations they perform and how they connect to each other. Ultimately, the model is a function that converts our inputs to the desired outputs.

   (d) What is the distinction between inference and learning for this task?

   Your answer: Inference is using the model with its current weights to make an educated guess about the correct output. Learning is making an educated guess and making sure the output is correct, then adjusting the model's weights to improve the output for future examples.

2. **[8 points]** K-Nearest Neighbors

K-Nearest Neighbors is an extension of the Nearest-Neighbor classification algorithm. Given a set of points with assigned labels, a new point is classified by considering the K points closest to it (according to some metric) and selecting the most common label among these points. One common metric to use for KNN is the squared euclidean distance, i.e.

$$d(x^{(1)}, x^{(2)}) = \|x^{(1)} - x^{(2)}\|_2^2 \tag{1}$$

For this problem, consider the following set of points in $\mathbb{R}^2$, each of which is assigned with a label $y \in \{1, 2\}$:

x_1	x_2	y
1	1	2
0.4	5.2	1
-2.8	-1.1	2
3.2	1.4	1
-1.3	3.2	1
-3	3.1	2

(a) Classify each of the following points using the Nearest Neighbor rule (i.e. $K = 1$) with the squared euclidean distance metric.

Your answer:

x_1	x_2	y
-2.6	6.6	1
1.4	1.6	2
-2.5	1.2	2

(b) Classify each of the following points using the 3-Nearest Neighbor rule with the squared euclidean distance metric.

Your answer:

x_1	x_2	y
-2.6	6.6	1
1.4	1.6	1
-2.5	1.2	2

(c) Given a dataset containing n points, what is the outcome of classifying any additional point using the n-Nearest Neighbors algorithm?

Your answer: Classifying an additional point will result in the new point taking the most common label of the n closest points and it will use some tie breaking method if multiple labels have the same number of occurrences within this subset.

(d) How many parameters are *learned* when applying K-nearest neighbors?

Your answer: No parameters are learned during KNN. KNN uses a fixed datasets and statically compute the classification for input points base on this dataset. KNN does not use the dataset to compute new parameters used to improve its ability to classify new inputs.