CS 6140: Machine Learning — Fall 2021— Paul Hand

HW 2 REVISED

Due: Wednesday September 29, 2021 at 2:30 PM Eastern time via Gradescope.

Names: [Put Your Name(s) Here]

You can submit this homework either by yourself or in a group of 2. You may consult any and all resources. You may submit your answers to this homework by directly editing this tex file (available on the course website) or by submitting a PDF of a Jupyter or Colab notebook. When you upload your solutions to Gradescope, make sure to tag each problem with the correct page.

Question 1. *In this problem, you will fit polynomials to one-dimensional data using linear regression.*

(a) Generate training data (x_i, y_i) for $i = 1 \ldots 8$ by $x_i \sim \text{Uniform}([0,1])$, and $y_i = f(x_i) + \varepsilon_i$, where $f(x) = 1 + 2x - 2x^2$ and $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$ and $\sigma = 0.1$. Plot the training data and the function f.

Response:

(b) In this problem, you will find the best fit degree d polynomial for the above data for each d between 0 and 7. Find it with least squares linear regression by minimizing the training mean squared error (MSE)

$$\min_\theta \frac{1}{2} \sum_{i=1}^{8} \left(y_i - \sum_{k=0}^{d} \theta_k x_i^k \right)^2 \tag{1}$$

using the Normal Equations. Use numpy.linalg.solve to solve the Normal Equations instead of computing a matrix inverse. On 8 separate plots, plot the data and the best fit degree-d polynomial.

Response:

(c) Plot the MSE with respect to the training data (training MSE) as a function of d. Which value of d provided the lowest training MSE?

Response:

(d) Generate a test set of 1000 data points sampled according to the same process as in part (a). Plot the MSE with respect to the test data (test MSE) as a function of d. Which value of d provided the lowest test MSE?

Response:

Question 2. *Linear regression using gradient descent and TensorFlow*

(a) Repeat 1b for $d = 1$ but solve (1) with gradient descent using TensorFlow or comparable software. Plot the data and resulting best fit line, and plot the training MSE versus epoch number. An epoch occurs when you iterate through the whole training dataset once. You may find Sections 3.1–3.3 in Mattman helpful for setting up your code.

Response: