



NORTHEASTERN UNIVERSITY, KHOURY COLLEGE OF COMPUTER SCIENCE

CS 6220 Data Mining — Assignment 5

Due: March 15, 2023(100 points)

YOUR NAME
YOUR GIT USERNAME
YOUR E-MAIL

Gradient Descent - Logistic Regression

Yann LeCun, the Director of Facebook Research and one of the fathers of deep learning neural networks, got his start with the MNIST Dataset, which is widely regarded as the "Hello World" of Computer Vision. The MNIST dataset can be downloaded at [LeCun's website](#). Logistic regression's cost function is typically binary cross-entropy.

$$\mathcal{L}(W, b) = \sum_i y_i \log h_{W,b}(\mathbf{x}) + (1 - y_i) \log (1 - h_{W,b}(\mathbf{x}))$$

where

$$h_{W,b}(\mathbf{x}) = \sigma(W^T \mathbf{x} + b)$$

and

$$\sigma(z) = \frac{1}{1 + \exp(-z)}$$

1. Prove that the derivative of the sigmoid function is $\nabla_z \sigma(z) = \frac{\partial \sigma(z)}{\partial z} = \sigma(z) (1 - \sigma(z))$.
2. Calculate the gradient of \mathcal{L} with respect to W .
3. Calculate the gradient of \mathcal{L} with respect to b .
4. Implement a functions `updateW` and `update_b` from your calculated gradients.