EE 660 **Homework 1 (Week 2)** Posted: Sat. 9/4/2021
Jenkins v0.8 Due: Mon., 9/13/2021, 11:59 PM PDT

=> **To turn in your Homework 1 solution,** please upload 2 files to the HW1 (Week 2) assignment dropbox in D2L, as follows.

(1) **a single pdf file of your solutions / answers to all the homework problems.** Please note:

(a) Your work can be hand written if you like; just scan it in, or take a picture with your smartphone and use a scan app to convert it to pdf; the result should look like a document-scanner result, not a photograph. Do not upload pictures in native format; the quality will be very senstive to lighting, and they are often not entirely readable.

(b) Please check the pdf for readability before uploading, and keep the file size reasonable (less than 5 MB).

(c) Be sure to include all answers to questions in the computer problem in this file.

(2) **a second pdf file that contains all your computer code for Problem 1.** This must be machine readable (not a scan, not a screenshot), and in a single file.

=> **Both files must be in pdf format.** No Word, .jpg, .txt, .py, zip or rar files. Thank you for cooperating; our grading methods depend on submissions as described above.

Reading: Murphy 8.1 – 8.3.1, inclusive (this was covered in Lecture 4.) Also, read 8.3.2 lightly (most of this you have had in a previous class).

1. *Comparison of loss functions* in logistic regression (log exponential loss), percetpron, and mean-squared error criterion functions, for classification.

(a) For logistic regression based on MLE, the loss function is (see objective function on Lecture 4 p. 19):

$$E_i^{(lr)} = \ln\left[1+\exp\left\{-\tilde{y}_i \underline{w}^T \underline{x}_i\right\}\right] .$$

Let $s_i = \tilde{y}_i \underline{w}^T \underline{x}_i$. Plot $E_i^{(lr)}$ *vs.* s_i , for $s_i = -10 \le s_i \le 10$.

(b) For 2-class linear perceptron learning (from EE 559), the objective function is:

$$J(\underline{w}) = -\sum_{i=1}^{N} \left[\!\left[\tilde{y}_i \underline{w}^T \underline{x}_i \le 0 \right]\!\right] \tilde{y}_i \underline{w}^T \underline{x}_i = \sum_{i=1}^{N} E_i^{(p)}$$

Give an expression for the loss function $E_i^{(p)}$ in terms of $s_i = \tilde{y}_i \underline{w}^T \underline{x}_i$. Plot $E_i^{(p)}$ *vs.* s_i , for $s_i = -10 \le s_i \le 10$.

(c) For the MSE objective function in a 2-class linear classification problem, the MSE can be written:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left[\underline{w}^T \underline{x}_i - b_i\right]^2 = \sum_{i=1}^{N} E_i^{(mse)} \ .$$

in which b_i is the target value for data point i. Let the target value be $b_i = \tilde{\underline{y}}_i \quad \forall i$.

Write the loss function $E_i^{(mse)}$ in terms of $s_i = \tilde{y}_i \underline{w}^T \underline{x}_i$. (**Hint:** first insert $\tilde{\underline{y}}_i$ into the above expression for MSE, in an appropriate place, where it has no affect on the MSE result.) Plot $E_i^{(mse)}$ *vs.* s_i , for $s_i = -10 \le s_i \le 10$.

(d) Compare the plots of (a), (b), and (c) above. Describe how these 3 loss functions contribute differently to the objective function. (For example, compare the loss functions for correctly classified data points that are near the decision boundary, and that are far from the decision boundary; likewise, compare the loss functions for incorrectly classified data points that are near the decision boundary, and that are far from the decision boundary.)

2. *Comparison of linear regression using least squares, ridge, and lasso.*

[This computer problem will be posted as soon as it is ready.]

3. *Estimating model variance (σ^2) in linear regression.* Murphy Exercise 7.4. **Hint:** Start from Murphy Eq. (7.8), and assume $\hat{\underline{w}}$ is given.

4. *Estimating w_0 separately in linear regression.* Murphy Exercise 7.5.