CS559 Homework#2
Due: 3/3/2020 11:59 PM

Homework #2 is an individual work: each student must submit their own work. Use of partial or entire solutions obtained from others or online is strictly prohibited. Electronic submission on Canvas is mandatory.

Logistic Regression I [10 pts]

By using the result $\frac{d\sigma}{da} = \sigma(1-\sigma)$ for the derivative of the logistic sigmoid, show that the derivative of the error function

$$E(\boldsymbol{w}) = -\ln p(\boldsymbol{t}|\boldsymbol{w}) = -\sum_{n=1}^{N}\{t_n \ln y_n + (1-t_n)\ln(1-y_n)\}$$

for the logistic regression model is given by

$$\nabla E(\boldsymbol{w}) = \sum_{n=1}^{N}(y_n - t_n)\boldsymbol{\phi}_n$$

where $y_n = \sigma(a_n)$ and $a_n = \boldsymbol{w}^T\boldsymbol{\phi}_n$.

Logistic Regression II and SVM I [40 pts]

Using the data file "winequality-red.csv", create a logistic regression model to predict quality of red wines according to the several features.

1) Exploratory Data Analysis:
Find how each feature is associated to the target 'quality'.
Explain how you would do on data to make the dataset as a binary classification problem.
2) Build the Logistic Regression function without using a python library. Here is a brief pseudo code.
   a. Initialize the weight and bias.
   b. A sigmoid function
   c. Forward/backward Propagation functions $z = x_1 w_1 + \cdots + x_n w_n + b$
   d. Updating Parameters Function
   e. Prediction Function
   f. Main Function - Logistic Regression
   g. Calculate the accuracy
3) Do the logistic regression using the python library 'sklearn'.
4) Use SVM to predict the quality of red wines.
5) Compare the results between 2,3, and 4. Explain the results.

SVM II [10 pts]

Consider the logistic regression model with a target variable $t \in \{-1,1\}$. If we define $p(t=1|y) = \sigma(y)$ where $y(\boldsymbol{x})$ is given by $y(\boldsymbol{x}) = \boldsymbol{w}^T\phi(\boldsymbol{x}) + b$, show that the negative log likelihood, with the addition of a quadratic regularization term, takes the form

$$\sum_{n=1}^{N} E_{LR}(y_n t_n) + \lambda||\boldsymbol{w}||^2$$

where $E_{LR}(yt) = \ln(1+\exp(-yt))$.

CS559 Homework#2
Due: 3/3/2020 11:59 PM

SVM III [10 pts]
Given 10 points in Table 1, along with their classes and their Lagrangian multipliers α_i, answer the following questions:
a) What is the equation of the SVM hyperplane $h(x)$? Draw the hyperplane with the 10 points.
b) What is the distance of x_6 from the hyperplane? Is this within the margin of the classifier?
c) Classify the point $z = (3,3)^T$ using $h(x)$ from above.

data	x_{i1}	x_{i2}	y_i	α_i
x_1	4	2.9	1	0.414
x_2	4	4	1	0
x_3	1	2.5	-1	0
x_4	2.5	1	-1	0.018
x_5	4.9	4.5	1	0
x_6	1.9	1.9	-1	0
x_7	3.5	4	1	0.018
x_8	0.5	1.5	-1	0
x_9	2	2.1	-1	0.414
x_{10}	4.5	2.5	1	0

Gaussian Process I [20 pts]
1. Implement Gaussian process regression for univariate variables y and x:

$$y \sim f(x) + \epsilon$$
$$f \sim GP(0, k)$$
$$\epsilon \sim N(0, \sigma^2)$$

Use a squared exponential kernel:

$$k(x, x') = \exp\left(-\frac{1}{2\lambda}(x - x')^2\right)$$

where λ is a parameter controlling the kernel width.
Draw some plots of curves sampled from just the prior distribution. Try a handful of different settings for λ to see what effect is has. You will need to add a small value to the diagonal of your covariance matrix to ensure it is invertible.

2. Test your Gaussian process regression with the following example. Generate synthetic data from the model:

$$y = x \sin(x) + \epsilon,$$
$$\epsilon \sim \mathcal{N}(0, \sigma^2).$$

Start by generating x values randomly uniform on the interval $[0, 2\pi]$. Then generate your y values using $\sigma = 0.5$ and a sample size of $n = 50$.

Plot (1) your raw data, (2) the true answer, (3) your estimated posterior mean function from Gaussian regression, and finally (4) the 95% confidence region for your posterior mean.

VIDEO PRESENTATION: Due on 2/28/2020 Friday 11:59 PM
Xianqing Zou - Gaussian Process
Ashish Desai - SVM