CSCI 416/516 Homework #2

DUE: November 01, 2024, at 11:59 pm

~~DUE: October 25, 2024, at 11:59 pm~~

CSCI 416/516: Each Problem begins with an allocation of points, represented as [u pts/g pts]. If you are registered in CSCI 416, you can receive up to u pts on this Problem; if you are registered in CSCI 516, you can receive up to g pts on this Problem. The last Problem is optional for undergraduates (CSCI 416) but required for graduates (CSCI 516). **Write down which session you are in / are you a graduate or undergraduate student**.

Optional/Extra Credit Problem(s) for Undergraduates: There is/are Problem(s) required for graduates but optional to undergraduates. Undergraduates can earn up to 10 credits from Problems 1-6 already, and can earn up to 12 credits from Problems 1-7. On the other hand, graduates can learn up to 10 credits from 1-7, so the maximum credits they can earn from attempting all the questions is 10 out of 10 (while for undergrads this number is 12 out of 10).

Submission: For all the problems excluding the multiple choice problem(s), you need to **show all your works, steps, and calculations** if applicable, or **your justification/expalantion to the answer(s) you provide**. You should submit a PDF to Blackboard with your answers that are recognizable/intelligible. Preferably, you should use LaTeX. Note that if you are attempting Problem 7, you should submit an additional PDF to Blackboard.

- **Problem 1 [2pts/2pts]: Support Vector Machine.**
  We want to maximize the margin between the cluster of positive samples and the cluster of negative samples using SVM. Suppose the support vectors in the cluster of positive samples fall on the vertical line $\boldsymbol{\theta}^\top \boldsymbol{x}_+ = 1$, and the support vectors in the cluster of negative samples fall on the vertical line $\boldsymbol{\theta}^\top \boldsymbol{x}_- = -1$, what is the formula, expressed in terms of $\boldsymbol{\theta}$, that we want to maximize? Show the steps on how you reached the conclusion.

- **Problem 2 [2pts/2pts]: Support Vector Machine.**
  The optimization objective of SVM is given as

$$\min_{\boldsymbol{\theta}} C \sum_{i=1}^{N} [y_i \text{cost}_1(\boldsymbol{\theta}^\top \boldsymbol{x}_i) + (1 - y_i)\text{cost}_0(\boldsymbol{\theta}^\top \boldsymbol{x}_i)] + \frac{1}{2}\sum_{j=1}^{d} \theta_j^2 \tag{1}$$

  where cost_0 and cost_1 are defined using the hinge loss. Explain the difference between the scenario in which the tunable hyperparameter C is large and the scenario in which C is small - what are we favoring, by making C large or small?

- **Problem 3 [2pts/1pts]: Support Vector Machine.**
How did we reach (for that $\sum_i a_j y_j = 0$, s.t. $\alpha_i \geq 0, \forall i$) Equation 2 from Equation 3?

$$\mathcal{J}(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle \tag{2}$$

$$\frac{1}{2} \sum_{j=1}^{d} \theta_j^2 - \sum_{i=1}^{n} \alpha_i (y_i (\boldsymbol{\theta}^\top x_i + b) - 1) \tag{3}$$

- **Problem 4 [1pt/1pt]: Kernels.**
Which of the following is not true about the kernel trick in SVMs?
  - A. The kernel function computes the dot product of two vectors in a transformed feature space.
  - B. The kernel function always transforms the data to a higher-dimensional space.
  - C. The choice of kernel affects the computational complexity of training an SVM.
  - D. The kernel function allows SVM to operate in a transformed feature space without explicitly calculating the transformation.

- **Problem 5 [1pt/1pt]: Lagrangian Multiplier.**
Suppose a sample in the training dataset has a Lagrangian multiplier being 0. What does this say about this sample?

- **Problem 6 [2pt/1pt]: Joint Entropy.**
Given the observation in Table , Suppose $X = \{\text{Raining, Not raining}\}$, $Y = \{\text{Cloudy, Not Cloudy}\}$, give the above Table. What is the joint entropy, $H(X, Y)$? Show your work on how the conclusion is reached.

	Cloudy	Not Cloudy
Raining	50/100	2/100
Not Raining	24/100	24/100

- **Problem 7 (Optional to Undergraduates) [2pt/2pt]: Anaconda.** Run the Tutorial on your local computer for the lecture topic *Support Vector Machines and Kernels*. You should submit a separate PDF document that contains the results of the execution of the code in the Tutorial.