Machine Learning Midterm

This TWO-SIDED exam is open book. You may bring in your homework, class notes and laptops text- books to help you. You will have 1 hour and 15 minutes. Write all answers in the blue books provided. Please make sure YOUR NAME and EID are on each of your blue books. Square brackets [] denote the points for a question. ANSWER ALL FOUR QUESTIONS FOR FULL CREDIT

1. **Entropy**

   The KL definition of Mutual Information is shown below.

$$\begin{aligned} \mathrm{I}[\mathbf{x},\mathbf{y}] &\equiv \mathrm{KL}(p(\mathbf{x},\mathbf{y})\|p(\mathbf{x})p(\mathbf{y})) \\ &= -\iint p(\mathbf{x},\mathbf{y}) \ln\left(\frac{p(\mathbf{x})p(\mathbf{y})}{p(\mathbf{x},\mathbf{y})}\right) \mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y} \end{aligned}$$

   [25]Use this definition to establish the relation showin this equation:

$$\mathrm{I}[\mathbf{x},\mathbf{y}] = \mathrm{H}[\mathbf{x}] - \mathrm{H}[\mathbf{x}|\mathbf{y}]$$

2. **Basic Probability**

   The Gamma probability distribution is defined as

$$Gamma(x|a,b) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}$$

   (a) [15] Compute $E(x)$ and $E(x^2)$

   (b) [10] Compute the variance

3. **Support Vector Machines** In the plane you are given **three** points, two classified +1 and one classifed -1.

   (a) [20] Show how you can use optimization to compute the best separation line between them.

   (b) [5] Make a sketch that illustrates your solution.

4. **Gibbs Sampling**

   In order to sample with a Markov chain, the chain has to exhibit **detailed balance**, given by

$$p^*(z)T(z,z') = p^*(z')T(z',z).$$

   In Gibbs sampling we sample the *ith* variable, keeping all the rest constant. thus our sample can be written as

$$Sample\ z_i^{(\tau+1)} \sim p(z_i|z_1^{(\tau+1)}, z_2^{(\tau+1)}, \ldots, z_{i-1}^{(\tau+1)}, z_{i+1}^{(\tau)}, \ldots, z_M^{(\tau)})$$

   which can be helpfully abbreviated as

$$Sample\ z_i^{(\tau+1)} \sim p(z_i|A)$$

   where A stands for the components that don't change for that sample with these notational simplifications

$$z = \{z_i, A\}, \tag{1}$$
$$z' = \{z_i', A'\}, \tag{2}$$
$$T(z,z') = p^*(z_i'|A) \tag{3}$$
$$p^*(z_i, A) = p^*(A)p(z_i|A) \tag{4}$$
$$\tag{5}$$

   (a) [5] What is the relation between A and A'?

   (b) [20] Use this notation to generate the steps that show that Gibbs sampling satisfies detailed balance.