Assignment 3

CMSC 726: Machine Learning
November 7^{th}, 2019 **Name:**

1. VC dimension of Hypothesis class: Consider the following hypothesis class $\mathcal{H} = \{h(x) : sign(\mathbf{w}^T\mathbf{x}) | x \in \mathbb{R}^n\}$, where $sign(z) = 1$ if $z \geq 0$ and $sign(z) = 0$ if $z \leq 0$. VCdim($\mathcal{H}$)

   (a) Show that VCdim $\geq n$. (**Hint 1**: Imaging a set of points $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}$ that correspond to the standard basis in $\mathbb{R}^n$, i.e. $\mathbf{x}_k^{(i)} = 1$ if $k = i$ and $\mathbf{x}_k^{(i)} = 0$ if $k \neq i$. What is the value of $\mathbf{w}$ that enables you to classify all points correctly using).

   (b) Show that VCdim $\leq n$. (**Hint 2**: Imagine that there exists a set of points $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n+1)}$ (more than n)such that they are shattered by $\mathcal{H}$. Form a matrix $\mathbf{H} = \mathbf{X}\mathbf{W}$ where $\mathbf{X} = [\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n+1)}]^T$ and $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_{2^n}]$. Here each $\mathbf{w}_i, 1 \leq i \leq 2^n$ corresponds to a possible labelling. Then prove that $rank(\mathbf{H}) \leq n$. This causes a contradiction in the assumption that $\mathcal{H}$ can shatter $n+1$ many points.)

2. (Programming Assignment) The Hoeffding inequality states that:

$$\mathbb{P}\left[\left|\frac{\theta_1 + \cdots + \theta_m}{m} - \mathbb{E}[\theta]\right| \geq \epsilon\right] \leq 2e^{\frac{-2m\epsilon^2}{(b-a)^2}}$$

   where θ_i's are generated in an i.i.d. fashion and each θ_i satisifies $a \leq \theta_i \leq b$.
   Let each θ_i be generated from $\mathbb{P}_\theta$ which is a uniform $[0,1]$ distribution.

   (a) Generate $k = 100$ many sets where each set S_i consists of $m = 100$ i.i.d. samples from $\mathbb{P}_\theta$.

   (b) What is the fraction of k sets that satisfies the following bound: $\mathbb{P}[|\frac{\theta_1+\cdots+\theta_m}{m} - \frac{1}{2}| \leq 0.1]$

   (c) Compare the number from part (b) with the Hoeffding bound.