IE643 Deep Learning : Theory and Practice **July-Dec 2025**

Quiz 1: Practice Questions

1 Practice Questions

1. Consider the training data $(x^t, y^t), t = 1, 2, \ldots$, where $x^t \in \mathbb{R}^{100}$, $\forall t = 1, 2, \ldots$. Denote the first 50 components of x^t as u^t and denote the last 50 components as p^t. Suppose $\max_{t=1,2,\ldots} \|u^t\|_1 \leq 1$ and $\max_{t=1,2,\ldots} \|p^t\|_\infty \leq 0.4$. Suggest a suitable least possible choice for the value of R in the mistake bound of perceptron.

2. Suppose you train a perceptron algorithm on a data set D_1. Design an algorithm to construct a new data set $D_2 \subseteq D_1$ such that when D_2 is used in training a perceptron afresh, it incurs the same number of mistakes as D_1. Is it possible that D_2 cannot be a proper subset of D_1? Explain. (Hint: Try with 2 dimensional data.)

3. Suppose $a_1x_1 + a_2x_2 = b$ is a separating hyperplane for a data set $D = (x^t, y^t),\ t = 1, 2, \ldots$, where $x^t = (x_1^t, x_2^t) \in \mathbb{R}^2$ and $y^t \in \{+1, -1\}$. Suppose b is non-zero, but we wish to make the hyperplane to pass through the origin. Explain what changes are to be made to the data set to achieve this purpose.

4. Give an example of a function $f : \mathbb{R} \to \mathbb{R}$ and a point $x^\star$ where $x^\star$ is neither local mimimum point nor local maximum point but $f'(x^\star) = 0$.

5. Consider a data set $D = \{(x^t, y^t)\},\ t = 1, 2, 3, \ldots$, where $x^t \in \mathbb{R}^d$ and $y^t \in \{+1, -1\},\ \forall t = 1, 2, 3 \ldots$. Recall that D is linearly separable if there exists a hyperplane $H = (w^*, b^*)$ (with $w^* \neq \mathbf{0}$) and $\gamma > 0$ such that $y^t(\langle w^*, x^t\rangle - b^*) > \gamma$ for every $t = 1, 2, 3, \ldots$. Assume that $\|w^*\|_2 = \beta$. Now consider a point x^p in D with $y^p = +1$. Find a point u on H closest to x^p and derive an expression for the distance between x^p and u. Explain how you constructed u.

6. In class, we discussed that the directional derivative of $f : \mathbb{R}^d \to \mathbb{R}$ (assumed to be continuously differentiable function) at a point $z \in \mathbb{R}^d$ along a non-zero direction v is of the form $f'(z; v)$ and this is related to the gradient of the function $\nabla f(z)$ at z as: $f'(z; v) = \langle \nabla f(z), v\rangle$. Prove this relationship.