5 Neural Networks [20 pts]

Let $f : [0,1] \times [0,1] \to \mathbb{R}$ be a 2-dimensional Lipschitz continuous function, that is $|f(x) - f(y)| \leq L\|x - y\|$, for all $x, y \in [0,1] \times [0,1]$. Let the $g(z)$ activation function be defined as $g(z) = 1$ if $z > 0$ and $g(z) = 0$ otherwise.

Knowing L and ϵ, construct a multi-layer neural network $\hat{f}(x)$ that has only linear and g activation functions, and $\hat{f}(x)$ is ϵ close to f, that is $|f(x) - \hat{f}(x)| \leq \epsilon$ for all $x \in [0,1] \times [0,1]$.

Hint: Partition $[0,1] \times [0,1]$. Remember, you did something similar to solve the "Neural Networks and Nearest Neighbors" question on homework 2: Question 3.2 part 1.

Partition the $[0,1] \times [0,1]$ domain of f into small boxes $B_1, \ldots, B_N$ of length $h = \epsilon/(L\sqrt{2})$, and let c_i be an arbitrary function value on box B_i. Since f is Lipschitz continuous, the function f will not change more than $|f(x) - f(y)| \leq L\|x - y\| \leq L\sqrt{2}h = \epsilon$ on these boxes. This way we can approximate $f(x)$ with a function $\hat{f}(x)$ that is constant (c_i) in each of these tiny boxes and $|f(x) - \hat{f}(x)| \leq \epsilon$ for all $x \in [0,1] \times [0,1]$.

Let $I_i(x)$ denote the indicator function of box B_i, that is $I_i(x) = 1$ if $x \in B_i$ and $I_i(x) = 0$ otherwise. Using this notation $\hat{f}(x) = \sum_{i=1}^{N} c_i I_i(x)$.

All that left is to prove that this piece-wise constant $\hat{f}(x) = \sum_{i=1}^{N} c_i I_i(x)$ can be represented with a neural network.

A box B_i is an intersection of four half spaces defined by the normal vectors of the four edges of the box. Let the equation of these half-spaces be denoted as $w_{i,j}^T x + b_{i,j} > 0$, for $j = 1, \ldots, 4$. We can say the $x \in B_i$, if and only if $w_{i,j}^T x + b_{i,j} > 0$ for all $j = 1, \ldots, 4$, that is $g(w_{i,j}^T x + b_{i,j}) = 1$ for all $j = 1, \ldots, 4$. In other words, the indicator function of B_i has the following equation:

$$I_i(x) = g(\sum_{j=1}^{4} g(w_{i,j}^T x + b_{i,j}) - 3.5)$$

Therefore

$$\hat{f}(x) = \sum_{i=1}^{N} c_i g(\sum_{j=1}^{4} g(w_{i,j}^T x + b_{i,j}) - 3.5)$$

This is what we wanted to prove.