EE488C Special Topics in EE <Deep Learning and AlphaGo> 1:00-3:00pm, October 24, 2016
EE, KAIST, Fall 2016

Mid-term exam

State your reasoning clearly. Show all necessary derivations. Try to simplify your final expression as much as possible. All your answers must be written in the boxes. If you need more space you can use the back of each page.

Name	Student ID	Total

1. (10 points) Find an example of two binary random variables X and Y such that X and Y are Bernoulli(1/2) and they are not independent.

2. (20 points) Consider a neural network given as follows.

$$y = \mathbf{w}^T \mathbf{h}, \ \mathbf{h} = g(\mathbf{W}^T \mathbf{x})$$

where $\mathbf{x} = (x_1, x_2)^T$ is the input vector, y is the output, $\mathbf{W} = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}$, $\mathbf{w} = (w_1, w_2)^T$, and $g(\cdot)$ is an elementwise ReLU. Assume $\mathbf{W} = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}$ and $\mathbf{w} = (\gamma, \gamma)^T$ and find conditions on α, β and γ with which this network can perform XOR, i.e., the output is 0 if the input is $(0, 0)^T$ or $(1, 1)^T$ and the output is 1 if the input is $(0, 1)^T$ or $(1, 0)^T$.

3. (20 points) Consider a problem of estimating θ from a random sample x generated according to Bernoulli(θ), where $0 \leq \theta \leq 1$. Let's consider a trivial estimator that always says α regardless of the sample it observes, where $0 \leq \alpha \leq 1$.

(a) Calculate the MSE of this estimator.

(b) Find the value of α that minimizes the worst-case MSE, i.e., find α that minimizes the following. What is the minimum worst-case MSE?

$$\max_{0 \leq \theta \leq 1} \mathbb{E}[(\alpha - \theta)^2]$$

4. (20 points) Let $J(x, y) = (x-2)^2 + 5(y-2)^2$ denote a cost function in terms of $(x, y) \in \mathbb{R}$. Let's add an L^1 regularization term $5\|(x, y)\|_1 = 5(|x| + |y|)$ and minimize $J(x, y) + 5(|x| + |y|)$. We can minimize this by following the steps below. The case $y \leq 0$ can be analyzed similarly, but the minimum of $J(x, y) + 5(|x| + |y|)$ is achieved when $y \geq 0$ and therefore we only need to analyze the following two cases to find the minimum value of $J(x, y) + 5(|x| + |y|)$.

(a) Assume $x \geq 0$ and $y \geq 0$ and minimize $J(x, y) + 5(|x| + |y|)$.

(b) Assume $x \leq 0$ and $y \geq 0$ and minimize $J(x, y) + 5(|x| + |y|)$.

5. (10 points) Consider the following one-tap IIR filter for averaging $u[k]$, where $0 < \alpha < 1$.

$$x[k] = \alpha x[k-1] + (1-\alpha)u[k]$$

Assume $u[k] = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases}$ and $x[k] = 0$ for $k < 0$. Calculate $\sum_{k=0}^{\infty} kx[k]$, which can be considered as an effective width of the impulse response.

6. (20 points) Design a continuous function $f : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\begin{aligned} f(x,0) &= 0,\ \forall x \in \mathbb{R} \\ f(0,y) &= 0,\ \forall y \in \mathbb{R} \\ f(1,1) &= 1 \\ f(-1,-1) &= 1 \end{aligned}$$

Try to make $f(x,y)$ as simple as you can such that it can generalize well for unseen samples.