Machine Learning Impt Stuff to note:

Saturday, October 27, 2018 8:01 PM

2015:

(1P) Multiple Choice: Neural Networks require feature mappings $\phi : x \mapsto \phi(x)$ in the lower layers to be designed by hand?

O Yes

O No

(2P) Suppose you have two data samples $x_1, x_2 \in \mathbb{R}^3$ with

$$x_i = \begin{pmatrix} x_i^{(1)} \\ x_i^{(2)} \\ x_i^{(3)} \end{pmatrix},$$

then write down a matrix transformation $A \in \mathbb{R}^{3\times3}$ such that the distance $\|Ax_1 - Ax_2\|$ between sample x_1 and x_2 is equal to

$$\|Ax_1 - Ax_2\| = \sqrt{(x_1^{(1)} - x_2^{(1)})^2 + 4(x_1^{(2)} - x_2^{(2)})^2}$$

Ans:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

[Hinge Loss]

1. given a set of $((f(x_1),y_1),\ldots,(f(x_N),y_N))$ consisting predictions $f(x_i)$ on features x_i and their corresponding ground truth labels y_i, write down the formula for hinge loss averaged over the number of samples N

   Ans:

   $$\frac{1}{N}\sum_{i=1}^{N} \max\{1 - y_i f(x_i), 0\}$$

[Mean Squared Error (MSE)]

1. given a set of $((f(x_1),y_1),\ldots,(f(x_N),y_N))$ consisting predictions $f(x_i)$ on features x_i and their corresponding ground truth labels y_i, write down the formula for mean squared error (MSE)averaged over the number of samples N

   Ans:

   $$\frac{1}{N}\sum_{i=1}^{N} (y_i - f(x_i))^2$$

EXTRA BONUS QNS:

Your score to this question, if you attempt it, will be given by

$$\max\{\,4-4x_1^2-8\alpha_1(1+x_2)-8\alpha_2(1-4x_1-x_2)-8\alpha_3(1-5x_1-x_2),\ \ 1\,\}$$

However, you are only allowed to choose the value of $x=(x_1,x_2)\in\mathbb{R}2$. I will be choosing the value of $\alpha=(\alpha_1,\alpha_2,\alpha_3)\in\mathbb{R}^3$ where each $\alpha_i\geq 0$. Write down your choice for x.

Note that if the student attempts the question, he/she will get at least 1 point!

The first term in the maximum function can be written as

$$4-8L(x,\alpha)$$

where $L(x,\alpha)$ is the Lagrangian of the one-dimensional SVM primal problem

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}x_1^2 \\ \text{subject to} \quad & (-1)(0\cdot x_1+x_2)\geq 1 \\ & (+1)(4\cdot x_1+x_2)\geq 1 \\ & (+1)(5\cdot x_1+x_2)\geq 1 \end{aligned}$$

The data set for this SVM problem is $\{(0,-1),(4,+1),(5,+1)\}$. Consequently, the primal problem achieves its optimal when $(0,-1)$ and $(4,+1)$ are support vectors. Thus, we need to solve

$$0\cdot x_1+x_2=-1, 4\cdot x_1+x_2=1$$

which gives $x_1=1/2$ and $x_2=-1$. If the student submits this choice for x, he/she will get 3 points.