CS282 Machine Learning: Quiz 2

CS282 TA Group
SIST
ShanghaiTech University

Problem 1

Given a data set $\mathcal{D} = \{(1, +1), (1.5, +1), (2, -1), (3, +1), (4, -1)\}$. Complete the basic set up of the learning problem:

1. Specify the hypothesis class $\mathcal{H}$, such that the function in $\mathcal{H}$ is continuous differentiable.
2. Choose a loss function ℓ for this problem.
3. Give a hypothesis $h \in \mathcal{H}$.

Solution. The answer is not unique, but you should be aware of that:

1. $\mathcal{H}$ is a set of hypothesis, usually determined by the parameters. For example, $\mathcal{H} = \{\boldsymbol{w}^T\boldsymbol{x} + b | \boldsymbol{w} \in \mathbb{R}^n, b \in \mathbb{R}\}$ in linear classifiers; $\mathcal{H} = \{k\text{NN} | k = 1, \ldots, n\}$ in kNN algorithms.
2. the loss function should be a **non-negative** function.
3. Choose an appropriate function $h \in \mathcal{H}$, it need not to be the optimal.

Remark Please refer to Section 1.1, 1.4 of *Learning From Data* for details.