

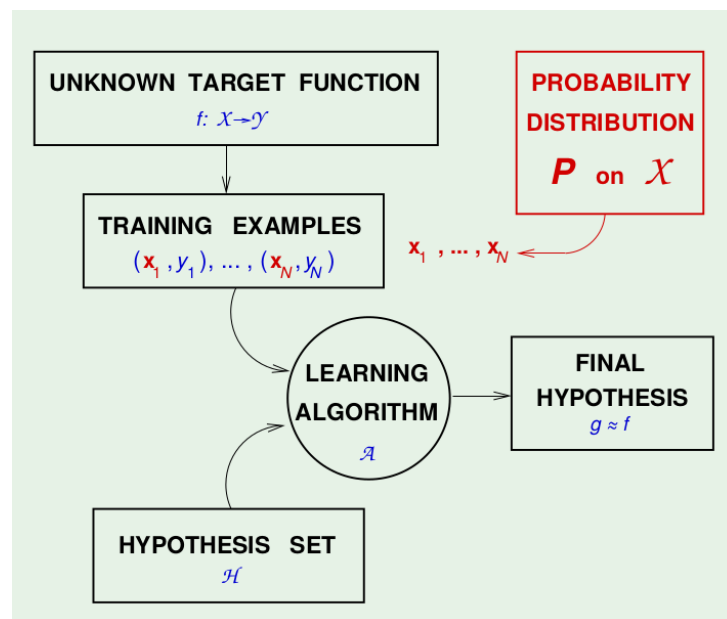
MAC0460/5832 – Introduction to Machine Learning

DCC / IME-USP — First semester/2021

Lista de exercícios 2 – Prazo para entrega: 07/06/2021

Esta lista cobre o conteúdo sobre a parte teórica de machine learning – *VC analysis* e *Bias-variance analysis* (*Lectures* 02, 05, 06, 07 e 08 da Caltech, ou capítulo 2 do livro-texto). Entregue uma versão escaneada/fotografada ou pdf com as respostas ou comentários. Note que para algumas questões não há, necessariamente, o certo ou não-certo para as respostas. A lista servirá para avaliar qual compreensão sobre o assunto cada um alcançou e também como guia para auto-avaliação e estudos adicionais individual.

1. Comente sobre o diagrama abaixo. O que o diagrama como um todo ilustra e o que cada componente representa?



2. O que é E_{in} e E_{out} ?
3. Quando consideramos a formulação teórica de aprendizado de máquina, uma das possibilidades é investigar o valor $|E_{in} - E_{out}|$. O que esse valor expressa e por que nos interessa investigar ele?

4. A desigualdade de Hoeffding, no contexto de aprendizado de máquina, com respeito a uma certa hipótese h , é dada por:

$$P\left(|E_{in}(h) - E_{out}(h)| > \epsilon\right) \leq 2e^{-2\epsilon^2 N}$$

Explique o significado dessa desigualdade.

5. A desigualdade de Hoeffding, no contexto de aprendizado de máquina, quando selecionamos uma hipótese de um espaço com M hipóteses é dada por:

$$P\left(|E_{in}(g) - E_{out}(g)| > \epsilon\right) \leq 2Me^{-2\epsilon^2 N}$$

Comente sobre a diferença entre essa desigualdade e a do item anterior.

6. O *bound* $2Me^{-2\epsilon^2 N}$ no item anterior foi obtido aplicando-se o *union-bound*. O que é *union-bound*?
7. O que são dicotomias ? O que é *growth-function* ? O que é *break point* ? Qual a relação entre eles ?
8. O que você entendeu sobre o processo envolvido na troca do M em $2Me^{-2\epsilon^2 N}$ pelo *growth-function* $m_{\mathcal{H}}(N)$? Qual o interesse em se fazer essa troca ? Qual é o novo *bound* obtido após a troca ?
9. Dissemos que a *VC dimension* relaciona-se com a expressividade do espaço de hipóteses. Comente sobre isso.
10. Como o *VC bound* é expresso em termos da *VC dimension* ?
11. Baseado no *VC bound*, explique como podemos calcular o número de amostras necessárias para se garantir uma certa precisão ϵ , com probabilidade $1 - \delta$, supondo que o espaço de hipóteses considerado tem dimensão VC igual a d_{VC} ?
12. Por que apenas garantir $|E_{in}(h) - E_{out}(h)| < \epsilon$ não é suficiente ?
13. Quais as similaridades e diferenças entre o *VC analysis* e o *Bias-variance analysis* ?
14. Escreva a sua opinião sobre quão úteis são os conteúdos cobertos nas *lectures* mencionadas para o entendimento sobre *Machine Learning*.