Statistics Learning Theory:VC-Dimension

2019.02.16

Motivation:Infinite-Size Classes

In the last week, we prove that finite classes are learnable. However, the infinite size classes may be learnable. Consider the following example.

Motivation:Example

Let $\mathcal{H}$ be the set of threshold functions over the real line, namely, $\{h_a : a \in R\}$, where $h_a(x) = I_{[x<a]}$. Then $\mathcal{H}$ is infinite, however it can be proved is learnable in the PAC model using the ERM algorithm.

VC-Dimension

The natural question arises: what is the sufficient conditions for learnability?
Answer: VC-dimension

VC-dimension

Definition (Restriction of $\mathcal{H}$ to C)

Let $\mathcal{H}$ be a class of functions from X to $\{0,1\}$ and let $C = \{c_1, \cdots, c_m\} \subset X$. The restriction of $\mathcal{H}$ to C is the set of functions from C to $\{0,1\}$ that can be derived from $\mathcal{H}$. That is,

$$\mathcal{H}_C = \{h(c_1), \cdots, h(c_m)) : h \in \mathcal{H}\}$$

Definition (Shattering)

A hypothesis class $\mathcal{H}$ shatters a finite set $C \subset X$ if the restriction of $\mathcal{H}$ to C is the set of all functions from C to $\{0,1\}$. That is, $|\mathcal{H}_C| = 2^{|C|}$

VC-dimension

Definition (VC-dimension)

The VC-dimension of a hypothesis class $\mathcal{H}$, denoted VCdim($\mathcal{H}$), is the maximalsize of a set $C \subset X$ that can be shattered by $\mathcal{H}$. If $\mathcal{H}$ can shatter sets of arbitraily large size we say that $\mathcal{H}$ has infinite VC-dimension.

VC-dimension:Examples

To show that $VCdim(\mathcal{H}) = d$ we need to show that

1. There exists a set C of size d that is shattered by $\mathcal{H}$
2. Every set C of size $d + 1$ is not shattered by $\mathcal{H}$

Examples: Threshold Functions

$C = \{c_1\}$, $\mathcal{H}$ shatters C, therefore, $VCdim(\mathcal{H}) \geq 1$. If an arbitrary set $C = \{c_1, c_2\}$ where $c_1 \leq c_2$, $\mathcal{H}$ does not shatter C.

Examples:Intervals

Let $\mathcal{H}$ be the class of intervals over R, namely,
$\mathcal{H} = \{h_{a,b} : a, b \in R, a < b\}$, where

The Fundamental Theorem of PAC learning

Sauer's Lemma and the Growth Function