1 The growth function

This is some section!

Question 1: Let $\mathcal{H}$ be any hypothesis set, where $h : \mathcal{X} \to \{-1, 1\}$ for all $h \in \mathcal{H}$. Abu-Mostafa 2015 defines that

$$\mathcal{H}(x_1, ..., x_N) = \{(h(x_1), ..., h(x_N)) \mid h \in \mathcal{H}\} \tag{1}$$

for specific $x_1, ..., x_N \in \mathcal{X}$, and further defines the growth function $m_{\mathcal{H}} : \mathbb{N} \to \mathbb{N}$ as

$$m_{\mathcal{H}}(N) = \max_{x_1,...,x_N \in \mathcal{X}} |\mathcal{H}(x_1, ..., x_N)| \tag{2}$$

Since $h(x) \in \{-1, 1\}$ for all $h \in \mathcal{H}$ and $x \in \mathcal{X}$, we must have for all $N \in \mathbb{N}$ and all $x_1, ..., x_N \in \mathcal{X}$:

$$\mathcal{H}(x_1, ..., x_N) \subset \{-1, 1\}^N = \{(y_1, ..., y_N) \mid y_1, ..., y_N \in \{-1, 1\}\} \tag{3}$$

which implies that for all $N \in \mathbb{N}$:

$$m_{\mathcal{H}}(N) \leq |\{-1, 1\}^N| = 2^N \tag{4}$$

This is true whether or not $\mathcal{H}$ is finite or infinite.

If we now assume that $\mathcal{H}$ is finite with $\mathcal{H} = \{h_1, ..., h_M\}$, then we get that for all $N \in \mathbb{N}$ and all $x_1, ..., x_N \in \mathcal{X}$:

$$\mathcal{H}(x_1, ..., x_N) = \{(h_i(x_1), ..., h_i(x_N)) \mid i \in \{1, ..., M\}\} \tag{5}$$

which implies that for all $N \in \mathbb{N}$ and all $x_1, ..., x_N \in \mathcal{X}$:

$$|\mathcal{H}(x_1, ..., x_N)| \leq M \tag{6}$$

which implies that for all $N \in \mathbb{N}$:

$$m_{\mathcal{H}}(N) \leq M \tag{7}$$

By line (1) and (7) we now have that for all $N \in \mathbb{N}$:

$$m_{\mathcal{H}}(N) \leq \min(M, 2^N) \tag{8}$$

Question 2: Let $\mathcal{H}$ be any hypothesis set, where $h : \mathcal{X} \to \{-1, 1\}$ for all $h \in \mathcal{H}$. Define the shattered sample sizes $\mathcal{N}_{\mathcal{H}}$ for $\mathcal{H}$ as:

$$\mathcal{N}_{\mathcal{H}} = \{N \in \mathbb{N} \mid m_{\mathcal{H}}(N) = 2^N\} \tag{9}$$

We can now rewrite Abu-Mostafa 2015's definition of the VC-dimension d_{VC} of $\mathcal{H}$ as

$$d_{VC}(\mathcal{H}) = \max \mathcal{N}_{\mathcal{H}} \tag{10}$$

where

$$\max \mathcal{N}_{\mathcal{H}} = \infty \tag{11}$$

if $\mathcal{N}_{\mathcal{H}} = \mathbb{N}$.

Assume that $\mathcal{H}$ is finite with $|\mathcal{H}| = M$. Let me now proof that for all $N \in \mathcal{N}_{\mathcal{H}}$:

$$N \leq \log_2 M \tag{12}$$

Assume that $N \in \mathcal{N}_{\mathcal{H}}$. By definition of $\mathcal{N}_{\mathcal{H}}$ we now have that

$$m_{\mathcal{H}}(N) = 2^N \tag{13}$$

Since $|\mathcal{H}| = M$, we know from question 1 that

$$m_{\mathcal{H}}(N) = \min(M, 2^N) \tag{14}$$

From combining line (13) and (14), we now get that

$$2^N \leq M \tag{15}$$

which implies that

$$N \leq \log_2 M \tag{16}$$

I have now proven that for all $N \in \mathcal{N}_{\mathcal{H}}$:

$$N \leq \log_2 M \tag{17}$$

This clearly means that

$$d_{VC}(\mathcal{H}) = \max \mathcal{N}_{\mathcal{H}} \leq \log_2 M \tag{18}$$

I have now proven that for all hypothesis sets $\mathcal{H}$, where $h : \mathcal{X} \to \{-1, 1\}$ for all $h \in \mathcal{H}$ and $|\mathcal{H}| = M$, we have that

$$d_{VC}(\mathcal{H}) \leq \log_2 M \tag{19}$$

Question 3: Let $\mathcal{H}$ be any hypothesis set, where $h : \mathcal{X} \to \{-1, 1\}$ for all $h \in \mathcal{H}$. As I argued in question 1, we have that for all $N \in \mathbb{N}$:

$$m_{\mathcal{H}}(N) \leq 2^N \tag{20}$$

which implies that for all $N \in \mathbb{N}$:

$$m_{\mathcal{H}}(2N) \leq 2^{2N} = \left(2^N\right)^2 \tag{21}$$

which together with line (20) implies that for all $N \in \mathbb{N}$:

$$m_{\mathcal{H}}(N) \leq (m_{\mathcal{H}}(N))^2 \tag{22}$$

Question 4: NOT SOLVED YET.

Question 5: Let $\mathcal{H}$ be any hypothesis set, where $h : \mathcal{X} \to \{-1, 1\}$ for all $h \in \mathcal{H}$. Theorem 2.4 in Abu-Mostafa 2015 states that if $m_{\mathcal{H}}(d) < 2^d$ for some $d \in \mathbb{N}$, then for all $N \in \mathbb{N}$:

$$m_{\mathcal{H}}(N) \leq \sum_{i=0}^{d-1} \binom{N}{i} \tag{23}$$

From this and question 4, it directly follows that if $m_{\mathcal{H}}(k) < 2^d$ for some $d \in \mathbb{N}$, then for all $N \in \mathbb{N}$:

$$m_{\mathcal{H}}(N) \leq N^{d-1} + 1 \tag{24}$$

Question 6: Let $\mathcal{H}$ be any hypothesis set, where $h : \mathcal{X} \to \{-1, 1\}$ for all $h \in \mathcal{H}$, and assume that there exists some $d \in \mathbb{N}$ such that $m_{\mathcal{H}}(d) < 2^d$. From question 1, we know that for all $N \in \mathbb{N}$:

$$m_{\mathcal{H}}(N) \leq 2^N \tag{25}$$

From question 6, we know that for all $N \in \mathbb{N}$:

$$m_{\mathcal{H}}(N) \leq N^{d-1} + 1 \tag{26}$$

Therefore, we get that for all $N \in \mathbb{N}$:

$$m_{\mathcal{H}}(N) \leq \min(N^{d-1} + 1, 2^N) \tag{27}$$

Let $\mathcal{N}_{\mathcal{H}}$ be defined as in question 2, and assume that $N \in \mathcal{N}_{\mathcal{H}}$. By definition of $\mathcal{N}_{\mathcal{H}}$ we now have that

$$m_{\mathcal{H}}(N) = 2^N \tag{28}$$

From combining line (XXXX) and (XXXX), we now get that

$$2^N \leq N^{d-1} + 1 \tag{29}$$

Define

$$A_d = \{N \mid 2^N \leq N^{d-1} + 1\} \tag{30}$$

Line (XXXX - XXXX) proofs that

$$\mathcal{N}_{\mathcal{H}} \subset A_d \tag{31}$$

which implies

$$d_{VC}(\mathcal{H}) = \max \mathcal{N}_{\mathcal{H}} \leq \max A_d \tag{32}$$

If the numerical value of A_d is of any interest, it can be found numerically for any d. A_d is namely finite for any $d \in \mathbb{N}$, because the lefthand inequality defining A_d is exponential, whereas the right hand side is polynomial.

Question 7: We see from the fact that the quantity

$$\binom{N}{d-1} \tag{33}$$

must be meaningful in order for line (XXXX) in question 5 to be meaningful that $d - 1 \leq N$. If this is not the case and $d - 1 > N$, we also have that

$$2^N \leq N^{d-1} + 1 \tag{34}$$

is true for all $N \in \mathbb{N}$, which implies that

$$A_d = \mathbb{N} \tag{35}$$

which implies that the bound on line (XXXX) turns into

$$d_{VC}(\mathcal{H}) \leq \infty \tag{36}$$

which is not a very useful bound.

2 VC-dimension

This is some section!

3 Airline Revisited

This is some section!

4 SVMs

This is some section!