# FTML practical session 13

14 juin 2025

TABLE DES MATIÈRES



## 1 BOUND ON THE ESTIMATION ERROR

**Step 2]**
The fact that

$$2 \sup_{h \in F} |R(h) - R_n(h)| \geqslant t \tag{1}$$

is equivalent to :

$$\cup_{h \in F} \Big( 2|R(h) - R_n(h)| \geqslant t \Big) \tag{2}$$

Then, Boole's inequality shows that :

$$P\Big( \cup_{h \in F} \Big( 2|R(h) - R_n(h)| \geqslant t \Big) \Big) \leqslant \sum_{h \in F} P\big(2|R(h) - R_n(h)| \geqslant t\big) \tag{3}$$

**Step 3]**
For each $h \in F$, we need to bound

$$P\big(2|R(h) - R_n(h)| \geqslant t\big) \tag{4}$$

We can apply Hoeffding's inequality to $X_i = l(h(x_i), y_i)$ :

— The $X_i$ are independent and identically distributed.

— $\forall i, E[X_i] = R(h)$ (see section 3.1.8 in lecture_notes.pdf)

Hence,

$$\forall h \in F, P\big(2|R(h) - R_n(h)| \geqslant t\big) \leqslant 2 \exp\Big( -\frac{nt^2}{2(b-a)^2} \Big) \tag{5}$$

The result can then be obtained by summing over the estimators $h \in F$.
**Step 4]**
We write

$$\delta = 2|F| \exp\Big( -\frac{nt^2}{2(b-a)^2} \Big) \tag{6}$$

Using **??**

$$\begin{aligned} P\Big(R(f_n) - R(f_a) < t\Big) &= 1 - P\Big(R(f_n) - R(f_a) \geqslant t\Big) \\ &\geqslant 1 - \delta \end{aligned} \tag{7}$$

Hence, with probability larger than $1-\delta$, $R(f_n) - R(f_a) < t$. We just need to express $t$ as a function of $\delta$.

$$\frac{\delta}{2|F|} = \exp\left(-\frac{nt^2}{2(b-a)^2}\right) \tag{8}$$

$$-2(b-a)^2 \log(\frac{\delta}{2|F|}) = nt^2 \tag{9}$$

$$2(b-a)^2 \log(\frac{2|F|}{\delta}) = nt^2 \tag{10}$$

$$2(b-a)^2 \left(\log(\frac{2}{\delta}) + \log(|F|)\right) = nt^2 \tag{11}$$

$$\sqrt{\frac{2(b-a)^2\left(\log(\frac{2}{\delta}) + \log(|F|)\right)}{n}} = t \tag{12}$$