



Additional 6-PACs



**Rianne de Heide; Ph.D. student@CWI
submission planned to AISTATS with
Kaufmann, Valko of INRIA Lille**



**Rémy Degenne; (former) postdoc@CWI;
NeurIPS 2019 paper with Koolen & Ménard**

Pierre Ménard; postdoc@INRIA Lille



Un-Expected Bernstein

- Learn a classifier $\hat{h} : \mathcal{X} \rightarrow \mathcal{Y} = \{0, 1\}$ from i.i.d. training data $Z^n = (X_1, Y_1), \dots, (X_n, Y_n) \sim P$
- **PAC-Bayes bound (McAllester '98)**: Error you make on future data bounded, with high probability, by error on training data plus complexity term:

$$\mathbf{E}_{(X,Y) \sim P}[\text{LOSS}(Y, \hat{h}_{|Z^n}(X))] \leq \frac{1}{n} \sum_{i=1}^n \text{LOSS}(Y_i, \hat{h}_{|Z^n}(X_i)) + \sqrt{\frac{\text{COMPLEXITY}(\hat{h}_{|Z^n})}{n}}$$

Un-Expected Bernstein

- Learn a classifier $\hat{h} : \mathcal{X} \rightarrow \mathcal{Y} = \{0, 1\}$ from i.i.d. training data $Z^n = (X_1, Y_1), \dots, (X_n, Y_n) \sim P$
- **PAC-Bayes bound (McAllester '98)**: Error you make on future data bounded, with high probability, by error on training data plus complexity term:

$$\mathbf{E}_{(X,Y) \sim P}[\text{LOSS}(Y, \hat{h}_{|Z^n}(X))] \leq \frac{1}{n} \sum_{i=1}^n \text{LOSS}(Y_i, \hat{h}_{|Z^n}(X_i)) + \sqrt{\frac{\text{COMPLEXITY}(\hat{h}_{|Z^n})}{n}}$$

- **WE:**

$$\mathbf{E}_{(X,Y) \sim P}[\text{LOSS}(Y, \hat{h}_{|Z^n}(X))] \leq \frac{1}{n} \sum_{i=1}^n \text{LOSS}(Y_i, \hat{h}_{|Z^n}(X_i)) + \text{STABILITY-COEFFICIENT}(\hat{h}_{|Z^n}) \cdot \frac{\text{COMPLEXITY}(\hat{h}_{|Z^n})}{n} + \sqrt{\frac{\text{CONST.}}{n}}$$