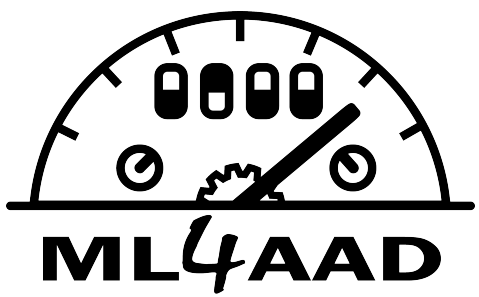
Machine Learning and Optimization for Algorithm Design

Feedback on Exercise 5

Marius Lindauer

University of Freiburg



Exercise 1

$$\mathrm{E}(\bar{x}) = \mathrm{E}(\tfrac{1}{n}\sum_{i=1}^{n} x_i) = \tfrac{1}{n}\sum_{i=1}^{n}\mathrm{E}(x_i) = \tfrac{1}{n}\cdot n\cdot\mu = \mu$$

$$\rightarrow \mu_{\bar{x}} = \mu$$

Exercise 1

$$\mathrm{E}(\bar{x}) = \mathrm{E}(\tfrac{1}{n}\sum_{i=1}^{n} x_i) = \tfrac{1}{n}\sum_{i=1}^{n}\mathrm{E}(x_i) = \tfrac{1}{n}\cdot n\cdot\mu = \mu$$

$$\rightarrow \mu_{\bar{x}} = \mu$$

$$Var(\bar{x}) = Var(\tfrac{1}{n}\sum_{i=1}^{n} x_i)$$

$$= \tfrac{1}{n^2} Var(\sum_{i=1}^{n} x_i)$$

$$= \tfrac{1}{n^2}\sum_{i=1}^{n} Var(x_i)$$

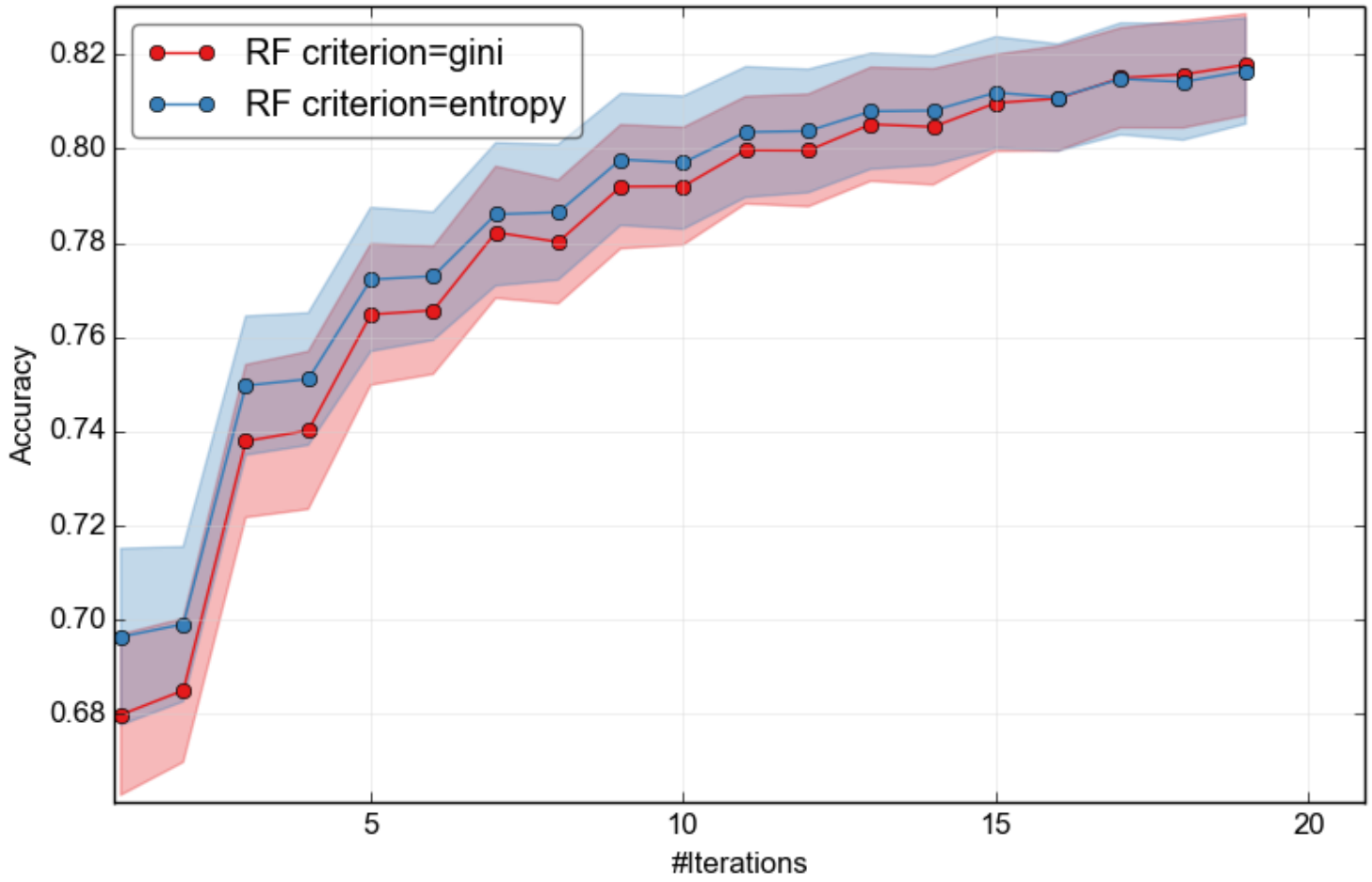
$$= \tfrac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{\sigma^2}{n}$$

$$\rightarrow \sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$$

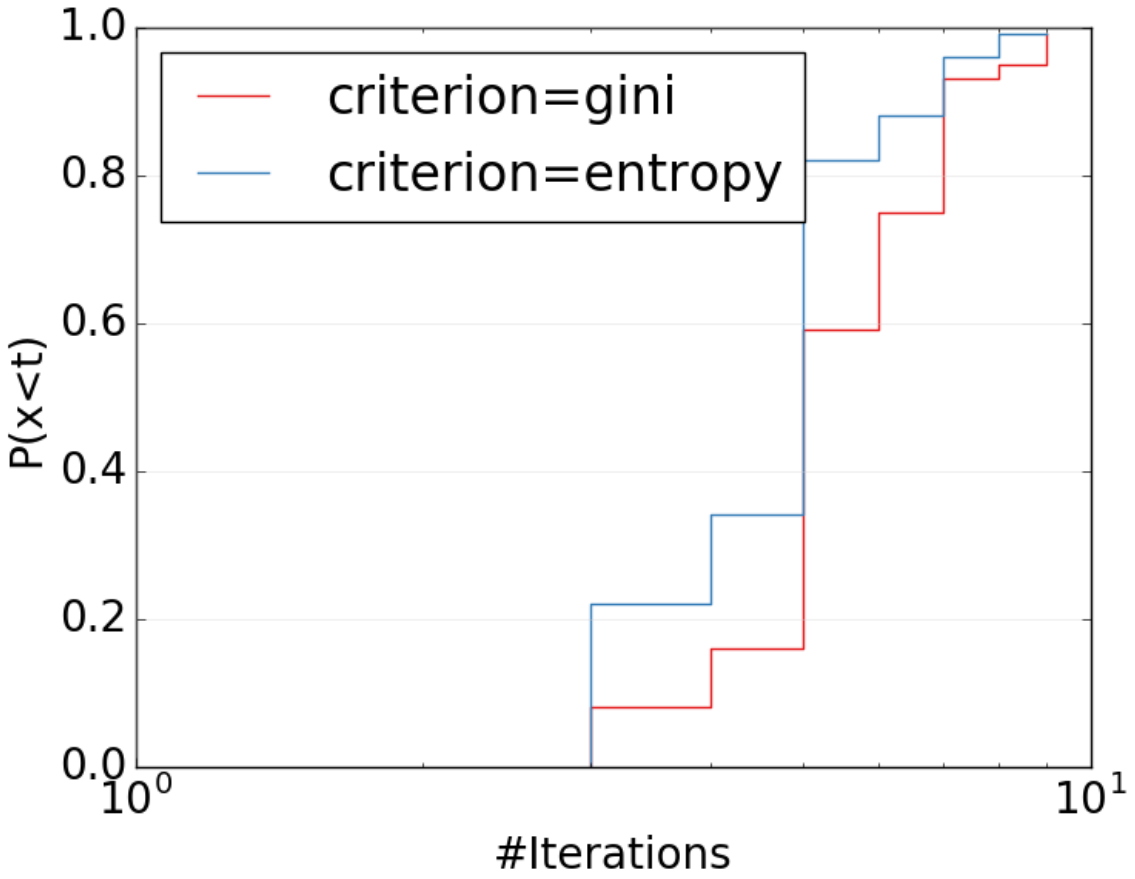
Exercise 2: Feedback

- Add labels on x-axis and y-axis.
- To compare two configurations, they should be shown in the same plot.
- Do not completely retrain the random forest for each new number iterations.
- Use the same colors in all plots.
- Empirical CDFs are step functions.

Exercise 2: Performance on Test



Exercise 2: RDF at acc=0.76



Exercise 3: Feedback

- Use log-scale on y-axis in box-plots (if appropriate).
- Something is wrong if all points are on the diagonal in a scatter plot.
- Use a paired permutation test when you have instances.

Score Overview