

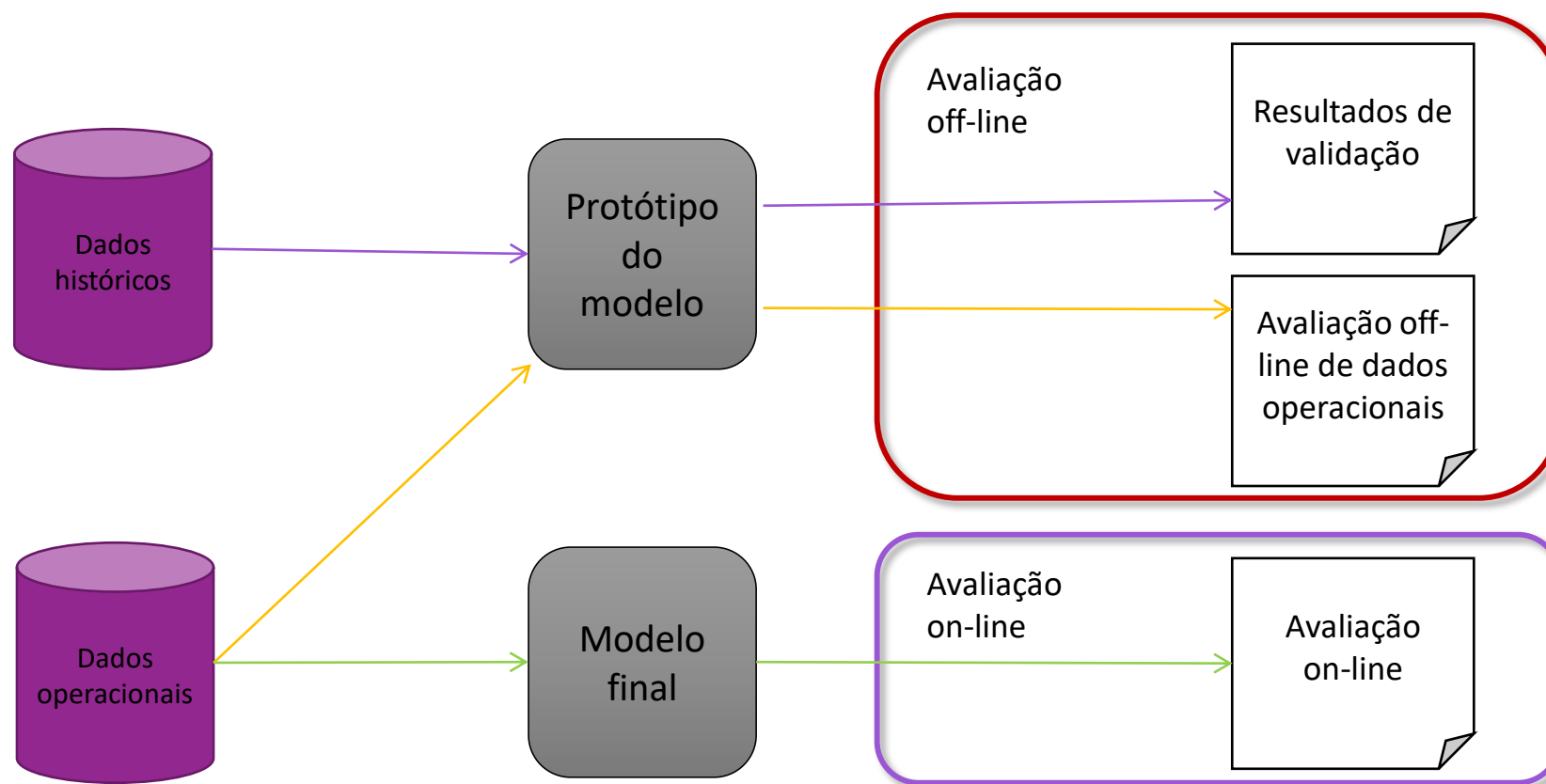
Avaliação de modelos

Prof. Hugo Bastos de Paula

Avaliação de modelos de Machine Learning

Cross-validation, RMSE, and grid search walk into a bar. The bar-tender looks up and says, “Who the heck are you?”

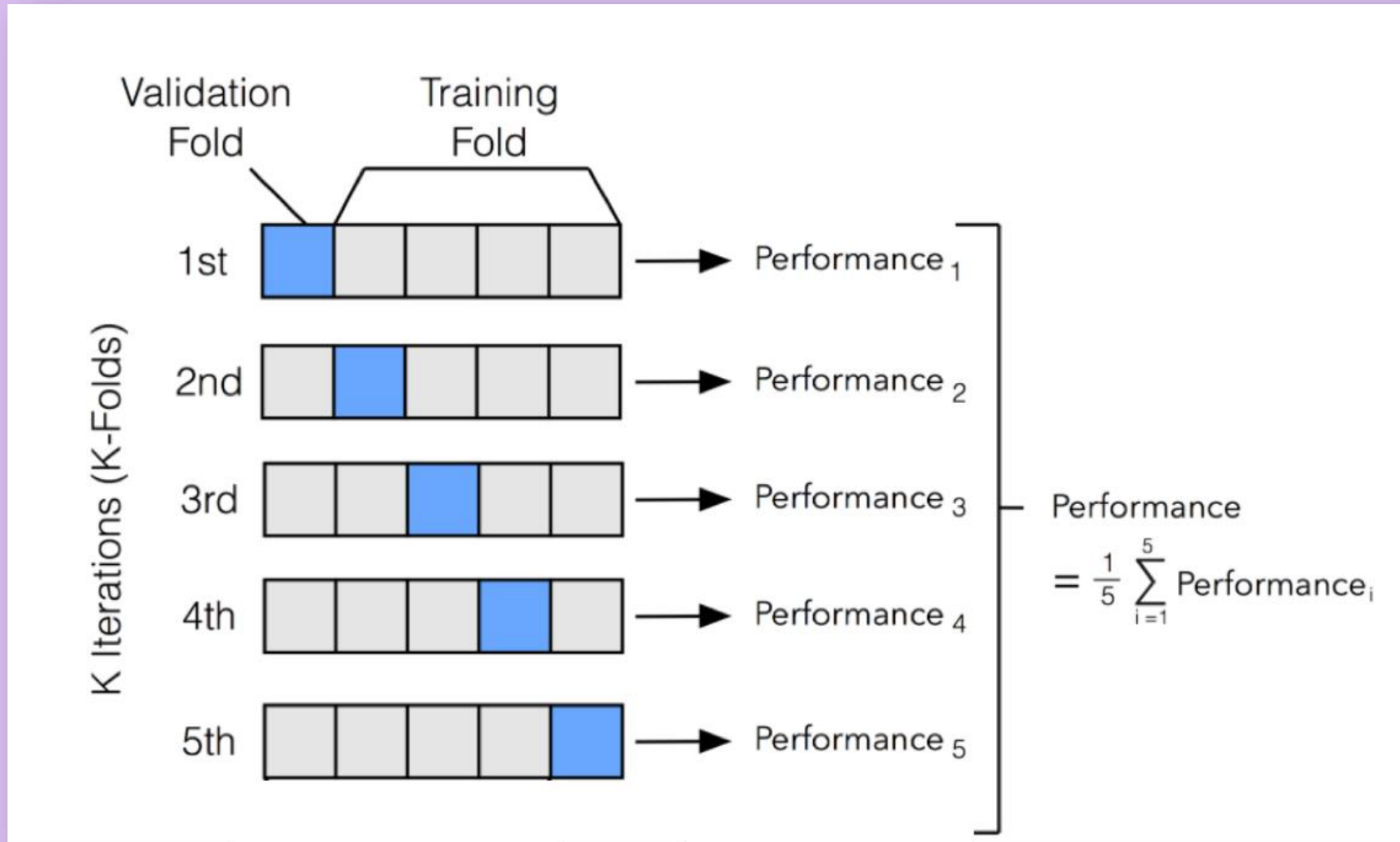
Alice Zheng (2015). Evaluating Machine Learning Models, A Beginner's Guide to Key Concepts and Pitfalls



Avaliação off-line

OBJETIVO: selecionar o melhor modelo para uma determinada tarefa.

Validação cruzada: cross validation

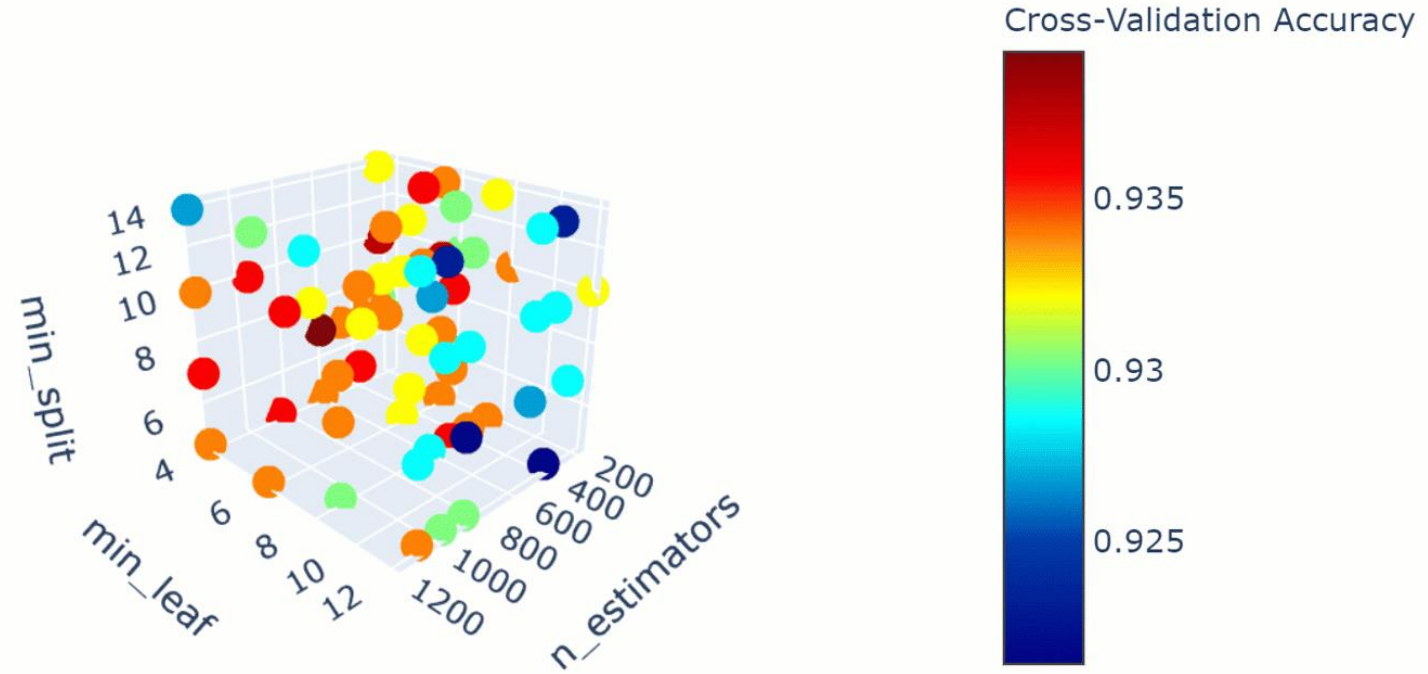


Parâmetros versus hiperparâmetros

Parâmetros de um modelo: variáveis ajustadas no processo de aprendizado.


Hiperparâmetros: precisam ser ajustados, mas não são aprendidos.

n_estimators 151.00 – 1200.00





Busca por hiperparâmetros

- Busca por hiperparâmetros ou *autotuning* são as técnicas para ajustar os hiperparâmetros de forma a maximizar a qualidade do modelo.
- 

Matriz de Confusão

	PREV. CLASSE A	PREV. CLASSE B	REVOCAÇÃO
CLASSE A	T_P	F_N	$T_P / (T_P + F_N)$
CLASSE B	F_P	T_N	
PRECISÃO	$T_P / (T_P + F_P)$		

F-measure ou F-score

$$F = 2 \times \frac{(PRECISAO \times REVOCAÇÃO)}{(PRECISAO + REVOCAÇÃO)}$$

Acurácia

- Porcentagem de elementos classificados corretamente (positivos ou negativos).

$$A = (T_P + T_N) / (T_P + T_N + F_P + F_N)$$

Acurácia média (acurácia por classe)

EXEMPLO: DETECÇÃO DE SPAM

	PREV. SPAM	PREV. NÃO SPAM
SPAM	80	20
NÃO SPAM	5	195

ACURÁCIA

$$A = \frac{80 + 195}{100 + 200} = 91,7\%$$

ACURÁCIA POR CLASSE

$$A_{SPAM} = \frac{80}{20+80} = 80\% \quad A_{NÃO\ SPAM} = \frac{195}{5+195} = 97,5\%$$

$$A = 80 + 97,5/2 = 88,75\%$$