

---

# Aprendizagem de Máquina

## Otimização de Hiperparâmetros

Telmo de Menezes e Silva Filho

[tmfilho@gmail.com](mailto:tmfilho@gmail.com)/[telmo@de.ufpb.br](mailto:telmo@de.ufpb.br)

[www.de.ufpb.br](http://www.de.ufpb.br)

UFPB



Departamento de  
**ESTATÍSTICA**

# Sumário

Introdução

Abordagens

Para Terminar

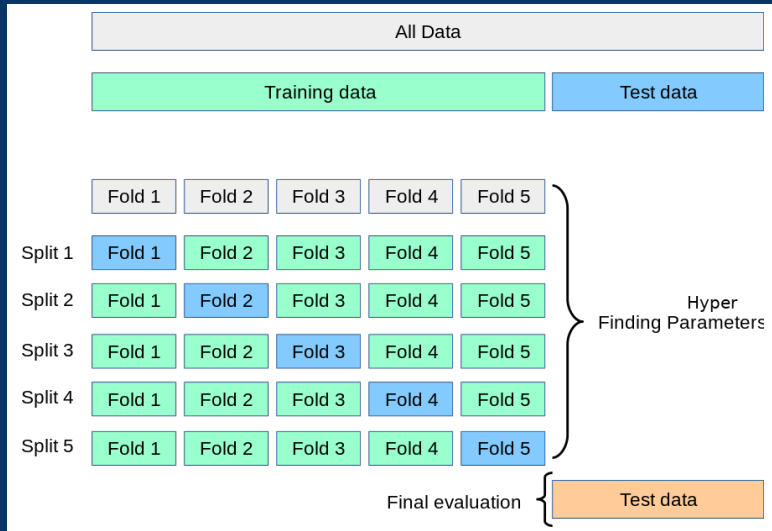


# Introdução

- ▶ Um hiperparâmetro é um parâmetro que controla o processo de aprendizagem
  - ▶ Note a diferença para os parâmetros que são aprendidos pelo ajuste do modelo, como pesos de redes neurais e coeficientes da regressão
- ▶ Como vimos ao longo do curso, cada algoritmo pode precisar de diferentes valores de hiperparâmetros que influenciam a capacidade de ajustar o modelo aos dados
- ▶ Assim, a otimização (*tuning*) de hiperparâmetros é um processo que busca encontrar a tupla de valores de hiperparâmetros que produzem o melhor modelo para um dado problema e uma dada métrica de avaliação
  - ▶ Costuma ser feita por meio de validação-cruzada



# Esquema de Tuning



# Grid Search

- ▶ A abordagem mais tradicional de tuning de hiperparâmetros é o grid search
- ▶ Trata-se de uma busca exaustiva através de todas as tuplas possíveis de hiperparâmetros dados valores aceitáveis para cada hiperparâmetro, fornecidos pelo usuário
- ▶ É necessário especificar uma métrica de performance
- ▶ Para parâmetros contínuos, é necessário definir limites e discretizar seus valores



# Grid Search

- ▶ Por exemplo, para um classificador SVM com kernel RBF (Gaussiano) precisamos otimizar pelo menos dois hiperparâmetros: o parâmetro de regularização  $C$  e a largura do kernel  $\gamma$
- ▶ Ambos são contínuos, então definimos um conjunto finito de valores razoáveis para nossa busca:

$$C \in \{10, 100, 1000\}$$

$$\gamma \in \{0.1, 0.2, 0.5, 1.0\}$$



# Grid Search

- ▶ O grid search vai então treinar um SVM com cada par  $(C, \gamma)$  possível e avaliá-lo no conjunto de validação
- ▶ Ao final do processo de validação-cruzada, a melhor configuração  $(C, \gamma)$  é retornada
- ▶ O processo todo é altamente paralelizável



# Random Search

- ▶ O random search evita a busca exaustiva por meio de uma seleção aleatória de combinações
- ▶ Isso pode ser aplicado aos conjuntos de valores discretizados do grid search, mas também pode ser aplicado a valores contínuos
  - ▶ É possível especificar a distribuição de probabilidade dos valores de cada hiperparâmetro
- ▶ o random search pode dar resultados melhores do que o grid search, principalmente quando dentre uma grande quantidade de hiperparâmetros, apenas alguns realmente influenciam o desempenho do algoritmo
  - ▶ E também por poder buscar valores fora da discretização necessária para o grid





# Algoritmos Evolutivos e Inteligência de Enxames

- ▶ A otimização populacional navega o espaço de hiperparâmetros usa algoritmos populacionais para navegar o espaço de hiperparâmetros de um algoritmo:
  1. Gere aleatoriamente uma população inicial de tuplas de hiperparâmetros
  2. Tome como fitness a métrica de desempenho escolhida, avaliada em uma validação-cruzada
  3. Aplique o algoritmo populacional escolhido até um critério de parada
- ▶ Essa abordagem tem sido muito usada para encontrar arquiteturas ideais de redes neurais profundas



# Para Terminar

- ▶ As diversas bibliotecas de AM costumam ter implementações de tuning de hiperparâmetros usando Grid Search ou Random Search
- ▶ Para garantir reprodutibilidade de resultados, é importante setar o random seed em todas as validações-cruzadas do processo
  - ▶ Isso facilita a paralelização do código
  - ▶ Com o random seed garantindo que os mesmos dados são usados, cada algoritmo pode ser tunado em um processo diferente do R ou do Python e depois os resultados podem ser avaliados
  - ▶ As bibliotecas frequentemente tem a opção de usar múltiplos processos (*workers*) e controlam isso automaticamente



# Sugestão de Atividade

- ▶ Trabalhe no projeto :D





# Aprendizagem de Máquina

## Otimização de Hiperparâmetros

Telmo de Menezes e Silva Filho

[tmfilho@gmail.com](mailto:tmfilho@gmail.com)/[telmo@de.ufpb.br](mailto:telmo@de.ufpb.br)

[www.de.ufpb.br](http://www.de.ufpb.br)

UFPB



Departamento de  
**ESTATÍSTICA**