Avaliação de Modelos

Prof. Jefferson T. Oliva

Aprendizado de Máquina e Reconhecimento de Padrões (AM28CP)
Engenharia de Computação
Departamento Acadêmico de Informática (Dainf)
Universidade Tecnológica Federal do Paraná (UTFPR)
Campus Pato Branco

Sumário

- Erro (Perda) em Problemas de Classificação e Regressão
- Decomposição de Viés-Variância do Erro Quadrático
- Classificação x Regressão
- Decomposição de Viés-Variância para a Perda 0-1
- Viés Estatístico x Viés Indutivo

Introdução

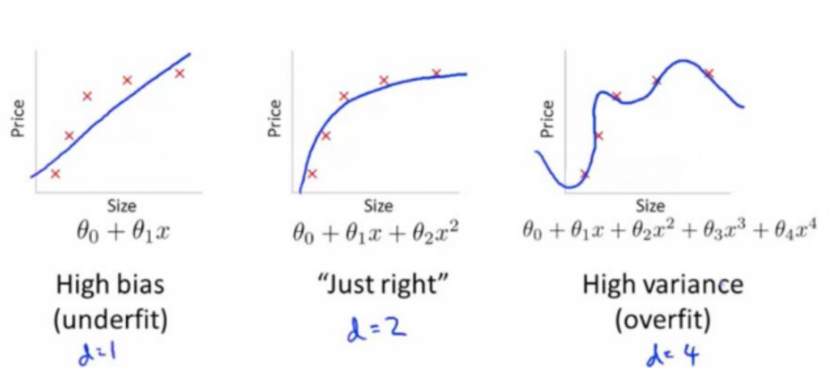
- A análise de desempenho de modelos depende do problema
  - Classificação: taxa de exemplos corretamente classificados
    - Acurácia, sensitividade, especificidade, f1-score, ...
  - Regressão: diferença entre os valor predito e o valor correto
    - Coeficiente de determinação (R^2), erro médio absoluto, erro médio quadrático, ...
- Média dos erros obtidos em diferentes execuções de um experimento
- O objetivo é errar o menos possível
  - Minimização da taxa de erro
  - Geralmente, não é possível medir a taxa de erro, ou seja, ela deve ser estimada

Erro (Perda) em Problemas de Classificação e Regressão

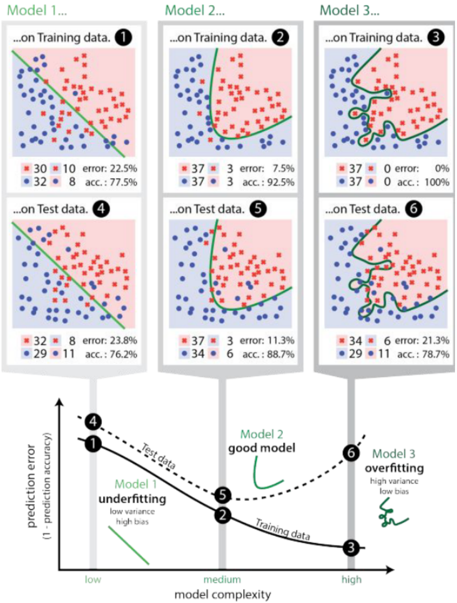
Erro (Perda) em Problemas de Classificação e Regressão

- É desejável que o modelo preditivo generalize "bem"
  - Alta acurácia de generalização
  - Baixo erro de generalização
- Suposição: os exemplos de treinamento e teste são independentes e distribuídos de forma idêntica (extraídos da mesma distribuição de probabilidade conjunta, $P(X, y)$)
- Para um qualquer modelo aleatório que não tenha sido ajustado ao conjunto de treinamento, esperamos que o erro de treinamento seja aproximadamente semelhante ao erro de teste
- O erro ou a acurácia do treinamento fornece uma estimativa com viés otimista do desempenho da generalização do modelo

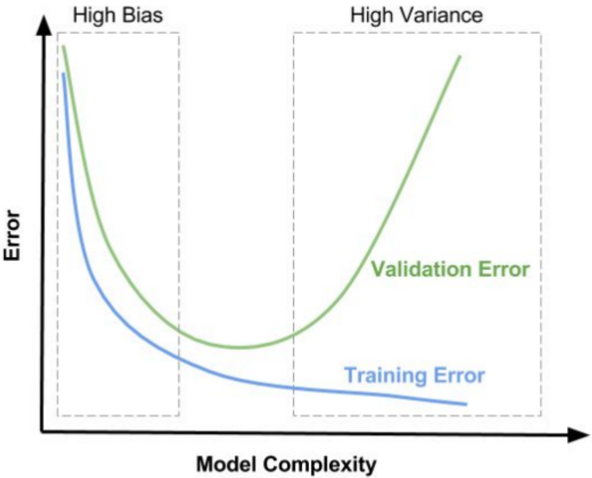
- Subajuste (*underfitting*): tanto o erro de treinamento quanto o de teste são altos
- Sobreajuste (*overfitting*): há diferença significativa entre o erro de treinamento e o de teste (onde o erro de teste é maior)

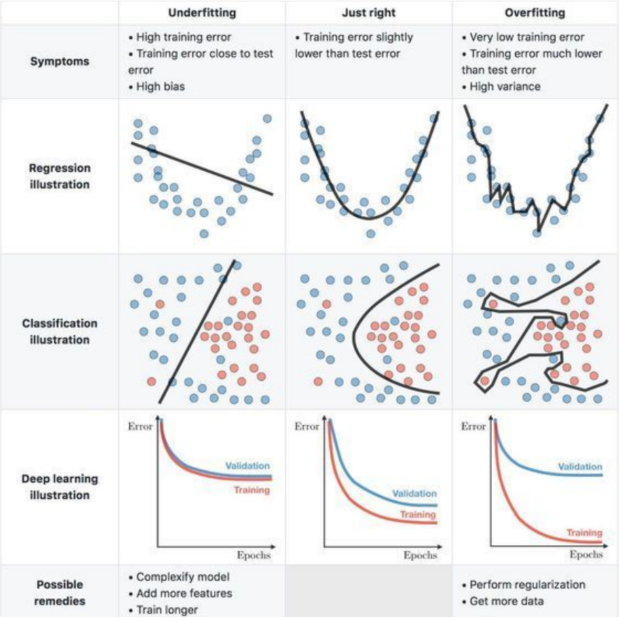


Erro (Perda) em Problemas de Classificação e Regressão



Erro (Perda) em Problemas de Classificação e Regressão





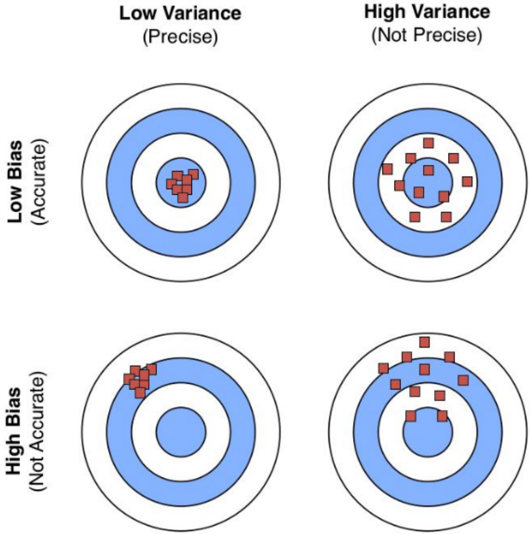
	Underfitting	Just right	Overfitting
Symptoms	• High training error • Training error close to test error • High bias	• Training error slightly lower than test error	• Very low training error • Training error much lower than test error • High variance
Regression illustration			
Classification illustration			
Deep learning illustration	Error Validation Training Epochs	Error Validation Training Epochs	Error Validation Training Epochs
Possible remedies	• Complexify model • Add more features • Train longer		• Perform regularization • Get more data

Decomposição de Viés-Variância do Erro Quadrático

Decomposição de Viés-Variância do Erro Quadrático

- A Decomposição da perda (*loss*) em relação ao viés e a variância é útil para a compreensão de algoritmos de aprendizado
  - Conceitos relacionados a subajuste e sobreajuste
- Ajuda a explicar por que métodos ensembles podem ter melhor desempenho do que modelos individuais
- Perda = Viés + Variância + Ruído

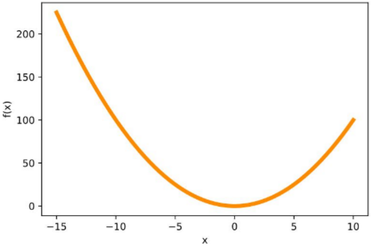
Decomposição de Viés-Variância do Erro Quadrático

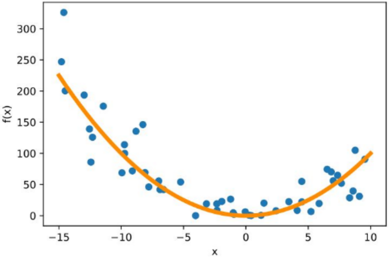


Decomposição de Viés-Variância do Erro Quadrático

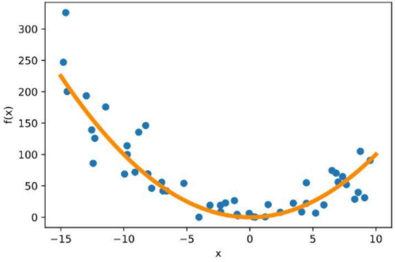
- $f(x)$: alguma função verdadeira (alvo)

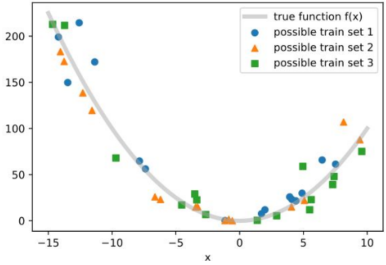
- Pontos azuis: conjunto de treinamento



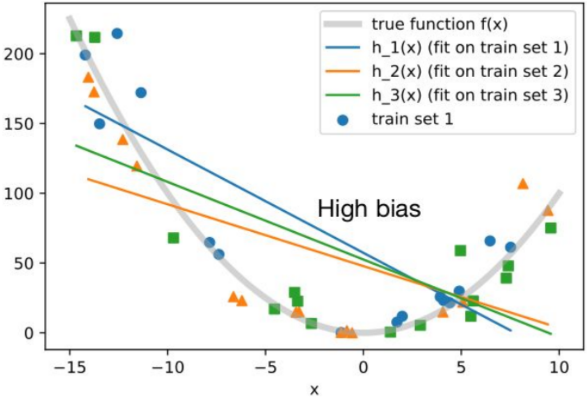


Decomposição de Viés-Variância do Erro Quadrático



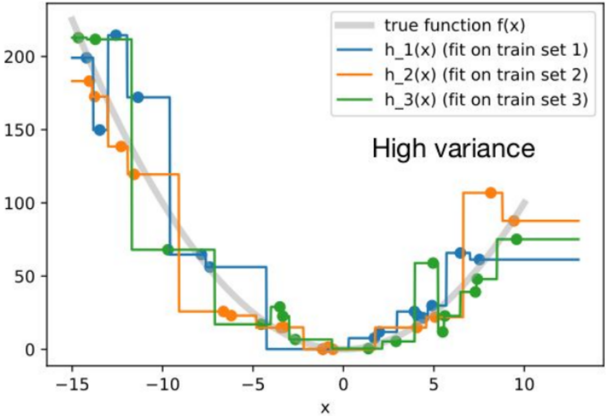


- Suponha que temos vários conjuntos de treinamento

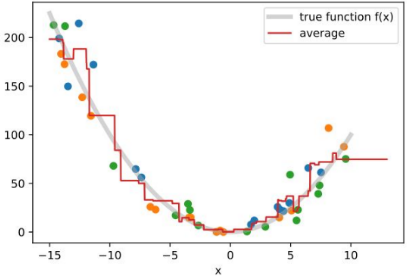


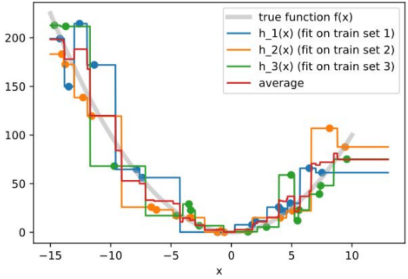
linear regression models

- Suponha que temos vários conjuntos de treinamento



- E se for calculada a média entre as predições?





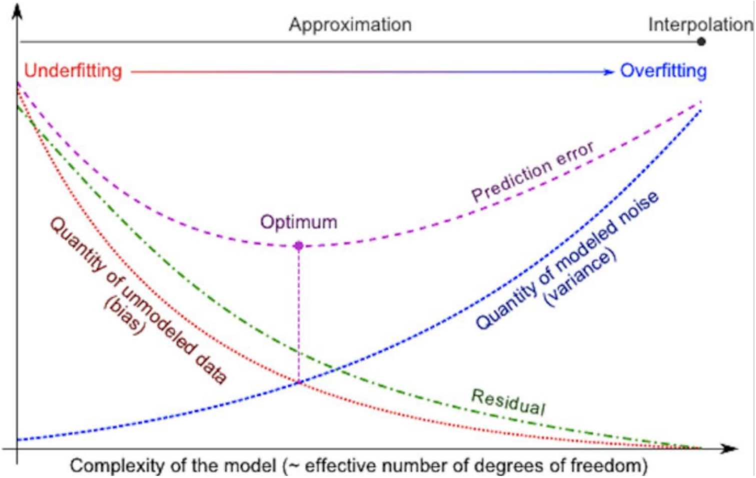
Decomposição de Viés-Variância do Erro Quadrático

- Dado um preditor $\hat{\theta}$ de um parâmetro θ
  - Poderia também ser uma função
    - Por exemplo, a hipótese é um preditor de uma função alvo

- Viés: diferença entre a média preditiva do modelo (treinado a partir de exemplos de treinamento) e os valores verdadeiros
  - O valor esperado está sobre o conjunto de treinamento

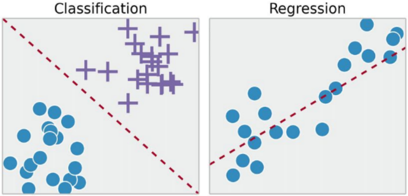
  - Viés $= E[\hat{\theta}] - \theta$

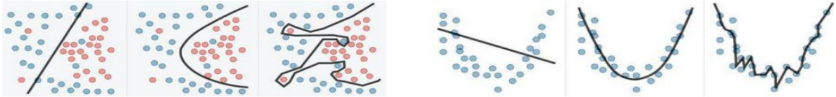
- Variância: fornece uma estimativa de quanto a predição varia à medida que variamos os dados de treinamento (*e.g*, reamostragem)
  - $Var[\hat{\theta}] = E[\hat{\theta}^2] - (\hat{\theta})^2 = E[(E[\hat{\theta}] - \hat{\theta})^2]$


Approximation
Interpolation
Underfitting
Overfitting
Prediction error
Optimum
Quantity of unmodeled data
(bias)
Quantity of modeled noise
(variance)
Residual
Complexity of the model (~ effective number of degrees of freedom)

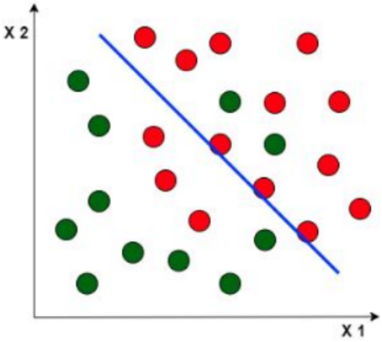
Classificação x Regressão

Classificação x Regressão

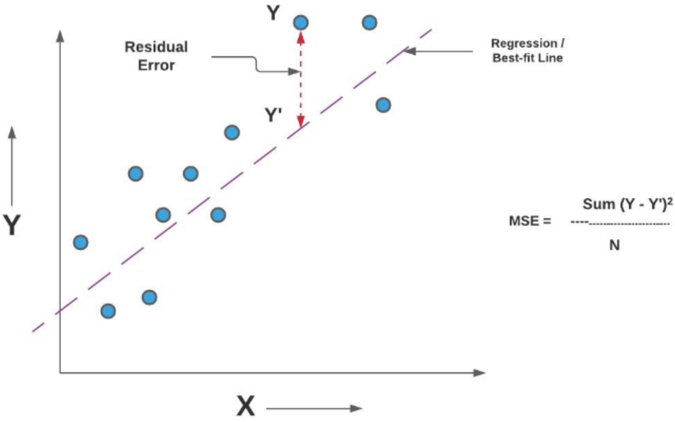




- Perda 0-1 (L_{0-1}) na classificação
  - Acurácia = 1 - taxa de erro
    - 1 - 0,2 = 0,8
  - $L_{0-1}(y_i, \hat{y}_i) = 1(y_i \neq \hat{y}_i)$
    - $L_{0-1} = 5$



- Perda por erro médio quadrático (*mean squared error* – MSE) na regressão



- Viés do ero quadrático
  - Viés $= E[\hat{\theta}] - \theta$
  - $Var[\hat{\theta}] = E[\hat{\theta}^2] - (\hat{\theta})^2 = E[(E[\hat{\theta}] - \hat{\theta})^2]$
  - Alvo: $y = f(x)$ (por simplicidade, ignoramos o termo ruído)
  - Predição: $\hat{y} = \hat{f}(x) = h(x)$
  - Erro quadrático: $S = (y - \hat{y})^2$

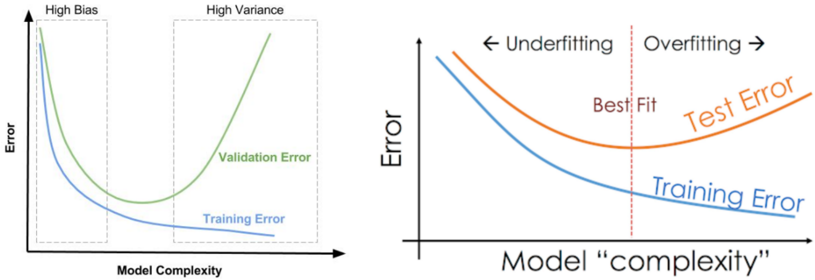
- Notação para a perda por erro quadrático
  - $(a-b)^2 = a^2 + b^2 - 2ab$

  - $S = (y - \hat{y})^2$
  - $S = (y - E[\hat{y}] + E[\hat{y}] - \hat{y})^2$
  - $S = (y - E[\hat{y}])^2 + (E[\hat{y}] - \hat{y})^2 + 2(y - E[\hat{y}])(E[\hat{y}] - \hat{y})$

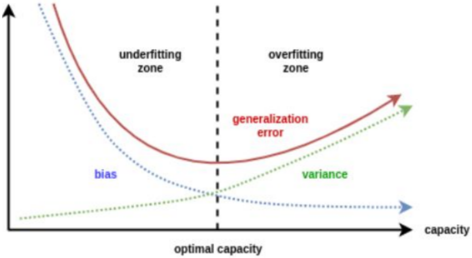
  - $E[S] = E[(y - \hat{y})^2]$
  - $E[(y - \hat{y})^2] = (y - E[\hat{y}])^2 + E[E[\hat{y}] - \hat{y}^2]$
  - $= \text{Viés}^2 + Var$

  - $E[2(y - E[\hat{y}])(E[\hat{y}] - \hat{y})] = 2E[(y - E[\hat{y}])(E[\hat{y}] - \hat{y})]$
  - $= 2(y - E[\hat{y}])E[E[\hat{y}] - \hat{y}]$
  - $= 2(y - E[\hat{y}])(E[E[\hat{y}]] - E[\hat{y}])$
  - $= 2(y - E[\hat{y}])(E[\hat{y}] - E[\hat{y}])$
  - $= 0$

- Como Viés^2 + *Var* está relacionado com subajuste e sobreajuste?

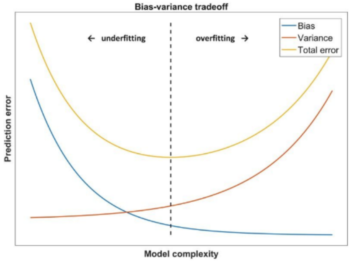


- Como Viés^2 + Var está relacionado com subajuste e sobreajuste?



Minimize 2 error sources!

bias-variance tradeoff...



Decomposição de Viés-Variância para a Perda 0-1

Decomposição de Viés-Variância para a Perda 0-1

- Diversas decomposições viés-variância relacionadas à Perda 0-1 foram propostas, mas todas apresentam deficiências consideráveis

- Perda quadrática:
  - $(y - \hat{y})^2$
  - Valor esperado sobre conjuntos de treinamentos para uma amostra específica
    $E[(y - \hat{y})^2] = (y - E[\hat{y}])^2 + E[(E[\hat{y}] - \hat{y})^2]$
    Viés^2 + Variância

- Predição principal $\rightarrow$ média $(E[\hat{y}])$

Decomposição de Viés-Variância para a Perda 0-1

- Perda 0-1
  - $L(y, \hat{y})$
  - $E[L(y, \hat{y})]$
- Predição principal $\rightarrow$ moda ($E[\hat{y}]$)
- Perda $= P(\hat{y} \neq y)$

$$\text{Viés} = \begin{cases} 1, & \text{se } y \neq \hat{y} \\ 0, & \text{caso contrário} \end{cases}$$

Viés Estatístico x Viés Indutivo

Viés Estatístico x Viés Indutivo

- Viés estatístico: diferença sistemática entre o valor esperado de um modelo e o valor verdadeiro (alvo)

$$\text{Viés} = E[\hat{\theta}] - \theta$$

- Viés indutivo: também denominado de viés de aprendizado de máquina, é a simplificação ou suposição incorreta sobre os dados
  - Determina o limite da capacidade do modelo em capturar a verdadeira relação entre as variáveis
  - O viés é o erro cometido por um modelo devido à sua flexibilidade insuficiente para representar a complexidade da função alvo

Viés Estatístico x Viés Indutivo

- A causa do viés estatístico pode ser dada por: escolha de métodos de predição inadequados, conjunto de dados não representativos, hiperparâmetros "incorretos" ...

- A causa do viés de aprendizado de máquina pode ser dada por: subajuste do modelo (*underfitting*), suposições inadequadas (*e.g.* usar regressão linear em dados não lineares), características insuficientes ...

- Exemplos de problemas reais relacionados ao viés
  - Na Amazon, o modelo de aprendizado de máquina para revisar currículos estava apresentando preferências em candidatos masculinos em relação ao feminino

- Exemplos de problemas reais relacionados ao viés
  - Em 2018, a polícia chinesa admitiu que difamou, por engano, uma mulher de negócios bilionária após o sistema de reconhecimento facial projetado para flagrar pedestres indisciplinados ter "capturado" ela em um anúncio de um ônibus

- Outros exemplos de problemas reais relacionados ao viés
  - Amazon's Rekognition
  - Predição de gênero
  - Base de dados de mamografia

- CASANOVA, D.
  Model evaluation. Aprendizado de Máquina.
  Slides. Engenharia de Computação. Dainf/UTFPR, 2020.

- DOMINGOS, Pedro. A unified bias-variance decomposition.
  In: Proceedings of 17th international conference on machine learning.
  Morgan Kaufmann Stanford, 2000. p. 231-238.

- RASCHKA, S.; MIRJALILI, V.
  Python Machine Learning.
  Packt, 2017.