

Problemas Clássicos da Computação

Fundamentos de Aprendizado de Máquinas

Felipe Augusto Lima Reis

felipe.reis@ifmg.edu.br



Sumário

- 1 Conceitos Básicos
- 2 Erros de Generalização
- 3 Treino, Validação e Teste
- 4 Métricas de Desempenho

CONCEITOS BÁSICOS

Introdução

- Algoritmos de aprendizado de máquinas tomam um conjunto de entradas e produzem saídas;
- Em um modelo de aprendizado básico, devemos nos preocupar, principalmente, com os seguintes tópicos:
 - Entrada de dados;
 - Saída predita pelo algoritmo;
 - Modelo de distribuição de dados;
 - Erros de classificação;
- Esses tópicos serão úteis no entendimento de outros conceitos relacionados à disciplina.

Modelo de Aprendizado Básico

- Relacionado às entradas e saídas dos algoritmos, temos a seguinte terminologia:
 - **Inputs**: correspondem às **entradas** x_i do algoritmo;
 - **Outputs**: correspondem às **saídas produzidas** \hat{y}_i pelo algoritmo;
 - **Target**: correspondem às **saídas esperadas** y_i , utilizadas em algoritmos supervisionados;
 - **Error**: corresponde à **inacurácia** E entre a saída produzida (*output*) e a saída esperada (*target*);
 - **Weights**: correspondem aos **pesos** w_{ij} , em conexões ponderadas entre dois nós i e j [Marsland, 2014].

[Marsland, 2014] utiliza a letra y para saídas produzidas (*outputs*) e t para saídas esperadas (*targets*).

Modelo de Aprendizado Básico

- Em relação à entrada para uma algoritmo de aprendizado, temos:
 - **Domain set:** conjunto de domínio \mathcal{X} , correspondente ao conjunto de objetos que serão rotulados;
 - **Label set:** conjunto de rótulos \mathcal{Y} para objetos do domínio;
 - **Dados de treinamento:** sequência finita de pares $\mathcal{X} \times \mathcal{Y}$, correspondente a pares rotulados;
- Em relação à saída do algoritmo, temos:
 - **Regra de aprendizado:** função $h : \mathcal{X} \rightarrow \mathcal{Y}$ utilizada para prever rótulos de novos pontos de dados
 - Também chamada de preditor, hipótese ou classificador [Shalev-Shwartz and Ben-David, 2014].

Modelo de Aprendizado Básico

- Em relação a um modelo de geração de dados, temos:
 - Instâncias (categorias) de dados são geradas com uma probabilidade de distribuição \mathcal{D} ;
 - Essa probabilidade \mathcal{D} atua sobre o conjunto de entradas \mathcal{X} ;
 - Uma função $f : \mathcal{X} \rightarrow \mathcal{Y}$ é utilizada para rotular cada uma das entradas \mathcal{X} ;
- Em relação à medida de sucesso de um algoritmo, temos:
 - Um erro de um classificador corresponde à probabilidade de que o algoritmo não preveja o rótulo correto para um ponto de dados aleatório;
 - O erro deve ser medido de acordo com a probabilidade de distribuição \mathcal{D} [Shalev-Shwartz and Ben-David, 2014].

Minimização de Risco Empírico¹

- Considere um conjunto de dados de treinamento \mathcal{S} , tomados aleatoriamente a partir de uma distribuição \mathcal{D} e rotulados com uma função f ;
- O objetivo de um problema de Aprendizado de Máquinas é encontrar um preditor $h_{\mathcal{S}}$ que minimize o erro em relação a \mathcal{D} e f (desconhecidos do algoritmo);
 - Espera-se que ao minimizar o erro desse conjunto amostral \mathcal{S} seja minimizado o erro de todo o espaço de dados;
 - Esse paradigma é chamado de **Minimização do Risco Empírico (ERM)** [Shalev-Shwartz and Ben-David, 2014];
- O aprendizado sobre um conjunto de dados amostrais, pode gerar um problema de generalização: *overfitting*.

¹ Empirical Risk Minimization (ERM).

ERROS DE GENERALIZAÇÃO

Introdução

- **Erro de generalização:** corresponde ao erro esperado de um preditor a respeito de dados ainda desconhecidos (não vistos)
 - Denotado também por perda (*loss*) ou erro verdadeiro²;
 - Medido de acordo com a probabilidade de distribuição \mathcal{D} ;
 - O erro é tido como esperado devido à inexistência de métrica para avaliar um conjunto desconhecido de informações [Shalev-Shwartz and Ben-David, 2014].

²Tradução literal de *true error*.

VIÉS E VARIÂNCIA

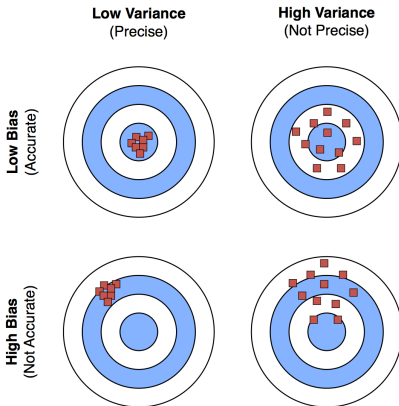
Viés e Variância

- Viés e variância descrevem a forma como modelos podem responder a um treinamento
 - **Viés:** descreve o quão bem o modelo corresponde ao conjunto de dados:
 - Alto viés indica que o modelo não consegue aprender e se aproximar do conjunto de dados;
 - Baixo viés indica que o modelo consegue representar adequadamente os dados de treinamento [Grant, 2019].

Viés e Variância

- Viés e variância descrevem a forma como modelos podem responder a um treinamento
 - **Variância**: corresponde à influência do subconjunto de dados de treinamento no aprendizado
 - Ao alterar o conjunto de treinamento, o algoritmo aprende de maneira diferente;
 - Alta variância indica que o modelo irá aprender de forma diferente a cada novo conjunto de dados;
 - Baixa variância indica que o modelo aprenderá características similares, independentemente do conjunto de treino [Grant, 2019].

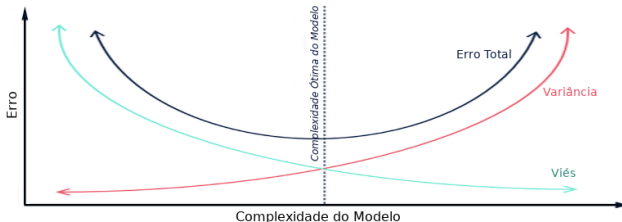
Viés e Variância



Fonte: [Raschka, 2018]

Viés e Variância

- *Trade-off* entre erros de viés e variância:
 - **Erro de viés:** corresponde ao uso de suposições erradas acerca do algoritmo, podendo causar perda informações relevantes e, consequentemente, subtreinamento (*underfitting*);
 - **Erro de variância:** corresponde ao erro de sensibilidade a pequenas flutuações no conjunto de treinamento, causando sobretreinamento (*overfitting*);



Fonte: Adaptado de [AI Pool, 2019]

UNDERFITTING E OVERFITTING

Underfitting



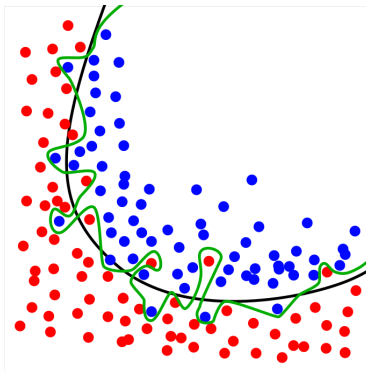
- **Underfitting**: situação onde um preditor possui desempenho ruim já no treinamento;
 - O modelo não consegue aprender a partir do conjunto de dados nem traçar relações entre variáveis;
 - Na prática, devido ao baixo desempenho na fase de treinamento, o modelo é descartado e nem mesmo é testado em um ambiente real;
 - Modelos com *underfitting* são geralmente modificados, para que possam melhorar sua representação.

Overfitting

- **Overfitting**: situação onde um preditor possui desempenho muito bom no treinamento e um muito ruim em situações reais [Shalev-Shwartz and Ben-David, 2014];
- Esse comportamento ocorre quando o preditor “decora” o conjunto de dados, devido ao excesso de treinamento;
- Segundo [Marsland, 2014], o tanto o sobretreinamento quanto o subtreinamento são danosos e perigosos aos algoritmos de aprendizado de máquinas
 - A falta de treino faz com que o algoritmo não aprenda suficientemente bem o problema;
 - O treino em excesso faz com que o algoritmo aprenda ruídos e inacurácias do conjunto de dados.

Overfitting - Exemplo

- Suponha dois conjuntos de dados (vermelhos e azuis);
- A linha verde representa um modelo sobretreinado, enquanto a linha preta representa um modelo treinado de forma adequada.



Fonte: [Icke, 2008]

Solução *overfitting* - ERM com viés indutivo

- Segundo [Shalev-Shwartz and Ben-David, 2014], são soluções para evitar o sobretreinamento:
 - ① Limitar o número de hipóteses possíveis
 - Para isso, limita-se o espaço de buscas;
 - Escolhe-se um conjunto de preditores antes da avaliação dos dados, com base em um conhecimento prévio do problema³;
 - ② Aumentar a quantidade de dados para treinamento
 - Ao aumentar a quantidade de dados para treinamento, os algoritmos terão maior dificuldade de aprender os dados;
 - O conjunto de dados é maior e, possivelmente, possui maior variância;
- Apesar de diminuir a possibilidade de *overfitting*, a técnica gera um *trade-off* de viés-variância.

³ Tal procedimento é chamado de viés indutivo (inductive bias), uma vez que adiciona um viés ao aprendizado.

CÁLCULO DE ERRO DE GENERALIZAÇÃO

Cálculo de Erro de Generalização

- O erro de generalização pode ser formalmente definido como:
[do Patrocínio Jr., 2018]

$$E_G(h, \mathcal{D}) = E_A(\mathcal{D}) + E_M + E_B$$

- onde:
 - $E_A(\mathcal{D})$: erro de aproximação, devido à variância;
 - E_M : erro de modelagem, devido ao viés;
 - E_B : erro Bayesiano
 - Erro decorrente de um conjunto de treinamento insuficiente ou inadequado para representação de características e separação das classes [do Patrocínio Jr., 2018].

Alguns autores definem o erro de generalização apenas como a soma entre erros de aproximação e modelagem.

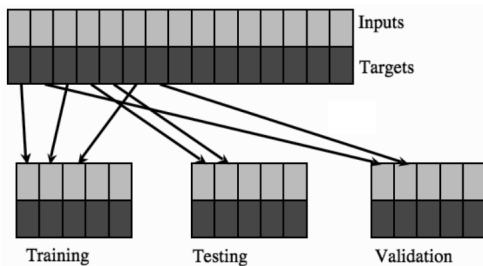
TREINO, VALIDAÇÃO E TESTE

Conjuntos de Treino, Validação e Teste

- Para treinamento de um algoritmo de aprendizado é apropriado utilizar 3 conjuntos: [Marsland, 2014]
 - **Conjunto de Treino:** destinado efetivamente ao treinamento do algoritmo;
 - **Conjunto de Validação:** destinado à avaliação de resultados durante a fase de treinamento;
 - **Conjunto de Testes:** destinado apenas à avaliação de resultados finais;
- **Importante:** O conjunto de testes deve ser utilizado somente após a conclusão de todo o treinamento
 - Em hipótese alguma, deve ser utilizado durante a fase de treinamento;
 - Deve, ainda, ser utilizado uma única vez.

Proporção entre conjuntos

- Considerando um conjunto único de dados, como dividi-lo em 3 subconjuntos?
 - Tipicamente utiliza-se a seguinte proporção 50:25:25, correspondente a treino, validação e testes;
 - Para casos em que há poucos dados, recomenda-se a divisão na proporção 60:20:20 [Marsland, 2014].



Fonte: [Marsland, 2014]

Modo de divisão dos conjuntos

- Não é apenas a quantidade de elementos em cada conjunto que importa: o modo de divisão também é importante;
- Algumas bases de dados colocam registros de classes de forma sequencial na base de dados:
 - Primeiros registros da base contêm *datapoints* da classe 1;
 - Em seguida, a base contém registros da classe 2, classe 3 e assim por diante;
 - Nesse caso, uma divisão sequencial de conjuntos causaria problemas no treinamento, uma vez que o algoritmo não aprenderia todas as classes;
- Uma solução possível seria o reordenamento aleatório dos conjuntos e nova divisão de conjuntos [Marsland, 2014]⁴.

⁴ Esse algoritmo é chamado de *Random Subsampling* [Ajitesh Kumar - DZone, 2018].

Modo de divisão dos conjuntos

- Em conjuntos muito pequenos, a utilização de um conjunto de validação pode impossibilitar o treinamento adequado;
 - Como existem poucos registros, a divisão pode comprometer a capacidade de generalização pela falta de exemplos;
 - Dentre as técnicas de divisão de conjuntos, destacam-se:
 - *k-Fold Cross validation*;
 - *Leave-one-out*;
 - *Repeated Cross validation*.

Base de Dados previamente divididas

- Algumas bases de dados já estão previamente divididas em conjuntos de treino, validação e testes
 - A divisão prévia tem como objetivo padronizar os testes, permitindo a comparação entre algoritmos;
 - Ex.: [BDS5 500 Dataset](#).
- Em outras bases, o conjunto de testes está indisponível para desenvolvedores;
 - Nela, a avaliação é feita em ferramentas da própria base, de modo a evitar que os algoritmos sejam treinados usando o conjunto de testes.
 - Ex.: [KITTI Road Dataset](#).

TÉCNICAS DE DIVISÃO DE CONJUNTOS

Técnicas de Divisão de Conjuntos

- A divisão da base de dados em conjuntos, de forma geral, busca possibilitar o treinamento e estimar o desempenho
 - Tal procedimento, ainda, deve ser feito sem perda da generalização e/ou adição de viés;
- Além do modelo indicado previamente, existem outras técnicas, como:
 - *Resubstitution*;
 - *Hold out*;
 - *k-Fold Cross validation*
 - *Leave-one-out*;
 - *Bootstrapping*.

Resubstitution

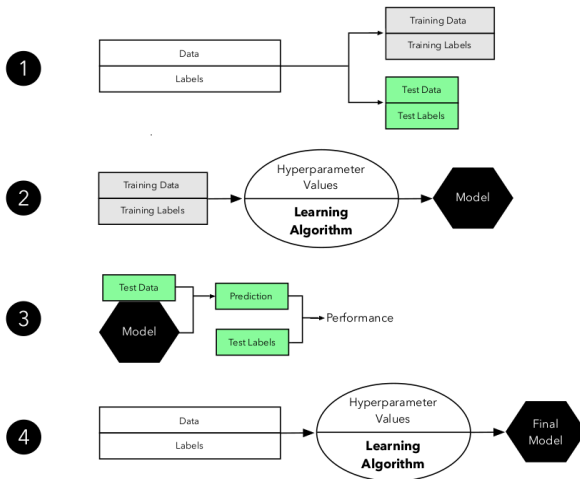
- Protocolo:
 - Todos os dados são usados para treinar e avaliar o modelo;
 - A taxa de erros é avaliada com base no mesmo conjunto de dados;
 - Considera-se que os dados “reais” são os mesmos do conjunto de treinamento (já conhecidos pelo algoritmo);
- Problema: extremamente otimista
 - Não é possível saber se o algoritmo conseguiu generalizar ou apenas memorizou os dados [Raschka, 2018];

Hold out



- Protocolo:
 - Divide-se um conjunto rotulado em apenas duas partes: treino e testes;
 - Ajusta-se o modelo aos dados de treinamento e são realizadas previsões dos rótulos do conjunto de teste;
 - A fração de previsões corretas é calculada com base nos rótulos verdadeiros do conjunto de teste [Raschka, 2018];
- Problema: o conjunto de testes é usado para verificação da melhor versão do treinamento
 - Tal procedimento pode gerar um falso nível de desempenho.

Hold out



Fonte: [Raschka, 2018]

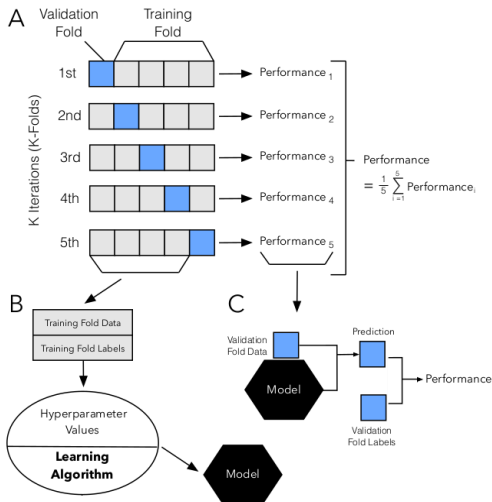
Repeated Hold out

- Variação do *Hold out*, na qual são feitas divisões aleatórias em conjuntos de treino e testes;
- A média das estimativas é usada para definição do desempenho do algoritmo;
- Proporções usuais: 90% para treino e 10% para testes [do Patrocínio Jr., 2018].

k-Fold Cross validation

- Destinado principalmente a conjuntos pequenos de dados;
- Também denominado *k-Fold Cross Validation*, *leave-some-out* ou *multi-fold cross-validation* [Marsland, 2014];
- Protocolo:
 - O conjunto de dados é dividido em k subconjuntos;
 - Um conjunto é usado como validação, enquanto o treinamento é feito nos demais;
 - Após uma primeira etapa, um subconjunto diferente é utilizado para validação enquanto outros são usados para treinamento;
 - O processo é repetido para todos os subconjuntos [Marsland, 2014];
- Nesse processo são treinados k diferentes modelos, ao invés de um só.

k-Fold Cross validation



Fonte: [Raschka, 2018]

Leave-one-out

- Caso extremo do método *k-Fold Cross validation*, no qual $k = m$, onde m é o número de exemplos
 - Nessa variação, um único registro é deixado de fora do treinamento [Shalev-Shwartz and Ben-David, 2014];
- Utilizado em conjuntos extremamente pequenos, onde não há possibilidade de novos dados;
- Também denominado pelos acrônimos LOO^5 ou $LOOCV^6$.

⁵ *Leave-one-out*.

⁶ *Leave-One-Out Cross-Validation*.

Repeated Cross validation

- Variação do método *k-Fold Cross validation*, onde os dados são divididos M vezes em k conjuntos;
 - O método de *cross validation* é executado em cada uma das divisões;
 - São geradas $M \times K$ estimativas;
 - A estimativa média é usada para definição do resultado final [do Patrocínio Jr., 2018].

Bootstrap

- Destinado principalmente a conjuntos pequenos de dados;
- Protocolo:
 - O conjunto de treinamento é selecionado aleatoriamente com repetição;
 - Os itens restantes, não selecionados, são usados para teste;
 - Um modelo é definido para cada uma das b amostras de bootstrap e são calculadas as taxas de precisão;
 - A taxa de erro do modelo é a média da taxa de erro de cada iteração [Ajitesh Kumar - DZone, 2018] [Raschka, 2018].

Bootstrap



Fonte: [Raschka, 2018]

MÉTRICAS DE DESEMPENHO

Métricas de Desempenho

- Além de avaliar a quantidade de dados necessários para testes do algoritmo, é necessário utilizar métricas que indiquem se o resultado é bom ou não [Marsland, 2014];
- Esta seção tem como objetivo descrever as principais métricas de avaliação de desempenho de algoritmos.

MÉTRICAS PARA CONJUNTOS DE DADOS BALANCEADOS

Matriz de Confusão

- A decisão tomada por um classificador pode ser representada por uma estrutura conhecida como matriz de confusão ou tabela de contingência.

		Valor Predito	
		Sim	Não
Real	Sim	Verdadeiro Positivo (TP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (TN)

Fonte: [Nogare, 2020]

		Outputs		
		C_1	C_2	C_3
C_1	5	1	0	
C_2	1	4	1	
C_3	2	0	4	

Fonte: [Marsland, 2014]

Matriz de Confusão

- Segundo [Davis and Goadrich, 2006], matrizes de confusão podem identificar 4 tipos de informação:
 - Verdadeiros Positivos (TP)
 - Valores corretamente classificados como positivos;
 - Falsos Negativos (FN)
 - Valores incorretamente classificados como negativos;
 - Falsos Positivos (FP)
 - Valores incorretamente classificados como positivos;
 - Verdadeiros Negativos (TN)
 - Valores corretamente classificados como negativos.

Métricas de Classificação Binária

- A partir da classificação dos resultados, podem ser geradas as seguintes métricas [Fawcett, 2006]:

- Taxa de Falsos Positivos (FPR)

- Quantidade de valores incorretamente classificados como positivos, FP , sobre o total de valores negativos, N ;

$$FPR = \frac{FP}{N}$$

- Taxa de Falsos Negativos (FNR)

- Quantidade de valores incorretamente classificados como negativos, FN , sobre o total de valores positivos, P .

$$FNR = \frac{FN}{P}$$

Métricas de Classificação Binária

- A partir das classificação dos resultados, podem ser geradas as seguinte métricas [Fawcett, 2006]:

- Taxa de Verdadeiros Positivos (TPR)

- Quantidade de valores corretamente classificados como positivos, TP , sobre o total de valores positivos, P ;

$$TPR = \frac{TP}{P}$$

- Taxa de Verdadeiros Negativos (TNR)

- Quantidade de valores corretamente classificados como negativos, TN , sobre o total de valores negativos, N .

$$TNR = \frac{TN}{N}$$

Acurácia

- Podemos considerar as seguintes métricas de acurácia:
 - Acurácia**
 - Quantidade de verdadeiros Positivos, TP , somado à quantidade verdadeiros negativos, TN , dividido pelo número de amostras [Fawcett, 2006];
- $$acurácia = \frac{TP + TN}{P + N}$$
- Taxa de Erro⁷**
 - Número de valores positivos e negativos classificados incorretamente sobre soma total de valores [Guo et al., 2012].

$$ER = \frac{FP + FN}{P + N}$$

⁷ Error Rate (ER).

Acurácia

- **Importante:** Para uso da métrica acurácia, devemos assumir, implicitamente, a existência da mesma quantidade de exemplos positivos e negativos
 - Esse conjunto de dados é conhecido como balanceado [Marsland, 2014].
- Ex.: Considere o algoritmo abaixo, para detecção de uma doença X, onde somente 1% dos indivíduos analisados estão efetivamente doentes
 - Esse algoritmo possui acurácia de 99%, no entanto, é incapaz de identificar um único paciente doente.

```
bool pacienteEstaDoente(){
    |   return False;
}
```

Fonte: Próprio autor

Precisão e Revocação

- Precisão⁸ e Revocação⁹ são métricas de classificação binária frequentemente usadas em Aprendizado de Máquinas;

- Precisão (*precision*)

- Mede a proporção de eventos de modelo que são verdadeiros;

$$precisão = \frac{TP}{TP + FP}$$

- Revocação (*recall*)

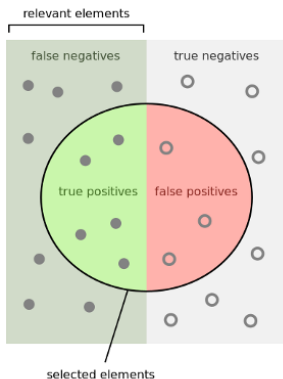
- Mede a proporção de eventos que ocorrem no domínio que são capturados pelos modelos [Fawcett, 2006].

$$revocação = \frac{TP}{TP + FN}$$

⁸ Denominado também por Valor Predito Positivo.

⁹ Denominado também por Sensibilidade.

Precisão e Revocação



How many selected items are relevant?

$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

How many relevant items are selected?

$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

Fonte: [Wikipedia contributors, 2020]

F-measure

- Para uma matriz de confusão simples, a *F-measure*¹⁰, é uma média harmônica entre a precisão e a revocação;
 - Busca balancear a importância de precisão e revocação;
 - Quando ambas possuem valores próximos, a métrica equivale a aproximadamente a média das duas medidas;

$$F\text{-measure} = \frac{2 \cdot \text{precisão} \cdot \text{revocação}}{\text{precisão} + \text{revocação}}$$

¹⁰ Também chamada de *F-score* ou *F₁ measure* [Marsland, 2014].

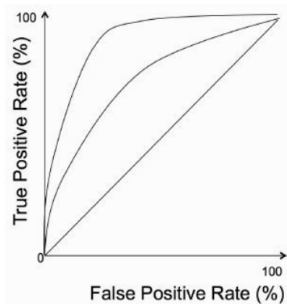
Curva ROC

- A curva **ROC** (*Receiver Operator Characteristic*) é uma representação gráfica que mostra o desempenho de um classificador binário [Fawcett, 2006];
 - Mostra a variação de amostras positivas classificadas corretamente em relação ao número de classificadas incorretamente;
 - Na curva, o eixo cartesiano y corresponde à taxa de verdadeiros positivos, enquanto o eixo x está relacionado à taxa de falsos positivos;
 - **Área sob a Curva**¹¹ corresponde à área sob a curva ROC, onde valores altos indicam boa precisão e revocação.

¹¹ Area Under Curve (AUC).

Curva ROC

- A figura abaixo exibe um exemplo de Curvas ROC.



Fonte: [Marsland, 2014]

Sensibilidade e Especificidade

- Relacionados à Curva ROC, é possível definir a **Sensibilidade** e a **Especificidade**

- **Sensibilidade** (*sensitivity*)

- Mesmo conceito de revocação;

$$\text{sensibilidade} = \frac{TP}{TP + FN}$$

- **Especificidade** (*specificity*)

- Número de verdadeiros negativos sobre a soma de falsos positivos e verdadeiros valores negativos [Marsland, 2014].

$$\text{especificidade} = \frac{TN}{FP + TN}$$

MÉTRICAS PARA CONJUNTOS DE DADOS DESBALANCEADOS

Acurácia Balanceada

- A acurácia balanceada é uma métrica para avaliação de classificadores binários, útil em bases de dados desbalanceadas;
- Pode ser definida como a média entre a sensibilidade e a especificidade [Marsland, 2014].

$$acurácia\ balanc. = \frac{sensibilidade + especificidade}{2}$$

Matthew's Correlation Coefficient

- Coeficiente utilizado em bases de dados desbalanceadas
 - Pode ser utilizado mesmo que as classes possuam tamanhos muito diferentes [Marsland, 2014] [Matthews, 1975];
- Retorna valores entre -1 e +1, que podem ser resumidos em:
 - +1: representa predições perfeitas;
 - 0: corresponde a predições aleatórias;
 - -1: discordância total entre predições e rótulos.

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

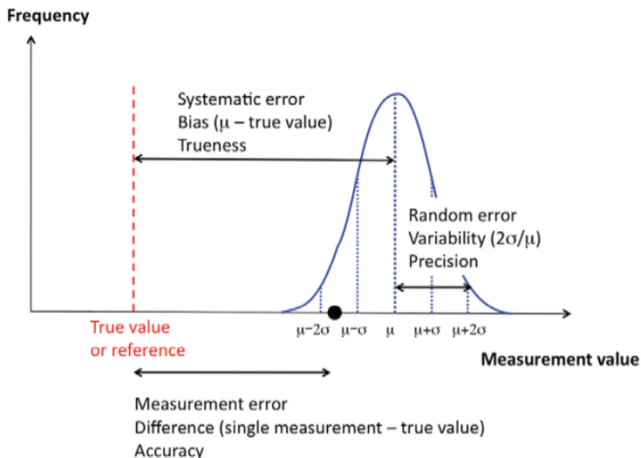
ACURÁCIA, PRECISÃO E VERACIDADE

Acurácia, Precisão e Veracidade

- Devido a problemas de tradução¹² e ao uso inadequado dos termos, as palavras acurácia e precisão são, frequentemente, consideradas sinônimos;
 - No entanto, elas são conceitualmente distintas;
- Definição:
 - **Acurácia**: definida como a proximidade das medidas em relação a um valor específico;
 - **Precisão**: corresponde ao quanto as medidas são similares entre si (proximidade entre as medidas);
 - **Veracidade** (*trueness*): distância média entre a saída correta e a previsão [Marsland, 2014].

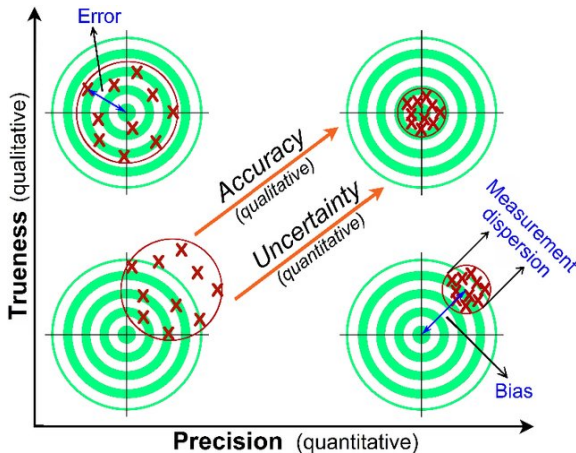
¹² A palavra “accuracy” é frequentemente traduzida como precisão.

Acurácia, Precisão e Veracidade



Fonte: [Squara et al., 2015]

Acurácia, Precisão e Veracidade



Fonte: [Villarraga-Gómez, 2016]

Referências I



AI Pool (2019).
Bias-variance tradeoff in machine learning.
Disponível em <https://ai-pool.com/a/s/bias-variance-tradeoff-in-machine-learning>.



Ajitesh Kumar - DZone (2018).
Machine learning: Validation techniques.
Disponível em <https://dzone.com/articles/machine-learning-validation-techniques>.



Davis, J. and Goadrich, M. (2006).
The relationship between precision-recall and roc curves.
In Proceedings of the 23rd International Conference on Machine Learning, ICML '06, pages 233–240, New York, NY, USA. ACM.



do Patrocínio Jr., Z. K. G. (2018).
Aprendizado de Máquina e Reconhecimento de Padrões - Conceitos Básicos / Teoria da Decisão.
Slides de Aula.



Fawcett, T. (2006).
An introduction to roc analysis.
Pattern Recognition Letters, 27(8):861 – 874.
ROC Analysis in Pattern Recognition.



Grant, P. (2019).
Introducing model bias and variance.
Disponível em <https://towardsdatascience.com/introducing-model-bias-and-variance-187c5c447793>.

Referências II



Guo, C., Mita, S., and McAllester, D. (2012).

Robust road detection and tracking in challenging scenarios based on markov random fields with unsupervised learning.

IEEE Transactions on Intelligent Transportation Systems, 13(3):1338–1354.



Icke, I. (2008).

Overfitting.

Disponível em <https://en.wikipedia.org/wiki/Overfitting>.



Kopec, D. (2019).

Classic Computer Science Problems in Python.

Manning Publications Co, 1 edition.



Marsland, S. (2014).

Machine Learning: An Algorithm Perspective.

CRC Press, 2 edition.

Disponível em: <https://homepages.ecs.vuw.ac.nz/~marsland/MLbook.html>.



Matthews, B. (1975).

Comparison of the predicted and observed secondary structure of t4 phage lysozyme.

Biochimica et Biophysica Acta (BBA) - Protein Structure, 405(2):442 – 451.



Nogare, D. (2020).

Performance de machine learning - matriz de confusão.

Disponível em <http://diegonogare.net/2020/04/performance-de-machine-learning-matriz-de-confusao/>.

Referências III



Raschka, S. (2018).

Model evaluation, model selection, and algorithm selection in machine learning.
CoRR, abs/1811.12808.



Richert, W. and Coelho, L. P. (2013).

Building Machine Learning Systems with Python.
Packt Publishing Ltd., 1 edition.



Shalev-Shwartz, S. and Ben-David, S. (2014).

Understanding Machine Learning: From Theory to Algorithms.
Cambridge University Press, 1 edition.

Disponível em: <http://www.cs.huji.ac.il/shaish/UnderstandingMachineLearning>.



Squara, P., Imhoff, M., and Cecconi, M. (2015).

Metrology in medicine: from measurements to decision, with specific reference to anesthesia and intensive care.

Anesthesia and analgesia, 120(1):66â75.



Villarraga-Gómez, H. (2016).

X-ray computed tomography for dimensional measurements.

In Digital Imaging 2016, An ASNT Topical Conference, pages 44–57. The American Society for Nondestructive Testing, Inc.



Wikipedia contributors (2020).

Precision and recall.

Disponível em https://en.wikipedia.org/wiki/Precision_and_recall.