

Universidade Federal do Pará
Programa de Pós-graduação em Engenharia de Processos
(PPGEP)
Disciplina: Tópicos Especiais em Engenharia de Processos
Prof. Dr. Alan de Souza
Aula 4

Setembro/2025

Na aula anterior...

1. Árvore de decisão como regressora;
2. Regressão logística para classificação;
3. Árvore de decisão de classificação;
4. Projetos práticos.

Métricas de avaliação de modelos de classificação

1. Matriz de confusão
2. Acurácia
3. Precisão
4. Recall (sensibilidade)
5. Especificidade
6. Fall-Out (Taxa de Falsos Positivos)
7. Medida F1
8. Curva ROC

Matriz de confusão

É uma tabela que resume os acertos e os erros de um modelo de classificação. Ela mostra, de forma organizada, onde o modelo acertou e errou.

		<i>Previsto</i>	
		Positivo	Negativo
<i>Real</i>	Positivo	Verdadeiro positivo (VP)	Falso Negativo (FN)
	Negativo	Falso Positivo (FP)	Verdadeiro Negativo (VN)

Matriz de confusão

Imagine uma linha de produção de peças metálicas.

O objetivo é usar um modelo de machine learning para classificar cada peça como "Defeituosa" ou "Correta".

		<i>Previsto</i>	
		Defeituosa	Correta
<i>Real</i>	Defeituosa		
	Correta		

Matriz de confusão

Depois de treinar o modelo, aplicamos em um lote de 100 peças. O que acontece?

		<i>Previsto</i>	
		Defeituosa	Correta
<i>Real</i>	Defeituosa	18	2
	Correta	5	75

- 18 peças ruins classificadas corretamente;
- 75 peças boas classificadas corretamente;
- 2 peças ruins que o modelo classificou como boas;
- 5 peças boas que o modelo classificou como ruins.

Matriz de confusão

- 18 peças ruins classificadas corretamente:
 - Modelo acertou.
- 75 peças boas classificadas corretamente:
 - Modelo acertou.
- 2 peças ruins que o modelo classificou como boas:
 - **Atenção: a fábrica mandou peças defeituosas para o cliente (perigoso).**
- 5 peças boas que o modelo classificou como ruins:
 - **Atenção: a fábrica descartou peças que estavam corretas (prejuízo).**

Matriz de confusão

- Resumindo: a matriz de confusão é como o “boletim” do modelo:
 - Mostra o que ele acertou,
 - Mostra onde ele se enganou,
 - E ajuda o especialista a decidir se é melhor reduzir falsos negativos (mais segurança do produto) ou reduzir falsos positivos (mais eficiência e menos desperdício).

DÚVIDAS?

Acurácia

		<i>Previsto</i>	
<i>Real</i>		Defeituosa	Correta
	Defeituosa	18 (VP)	2 (FN)
	Correta	5 (FP)	75 (VN)

Total: 100 peças.

- É a proporção de acertos totais
- É a média de acertos da classificação:

$$Acurácia = \frac{VP + VN}{Total}$$

$$Acurácia = \frac{18 + 75}{100} = \frac{93}{100} = 0,93 = 93\%$$

Precisão

Real	Previsto	
	Defeituosa	Correta
	Defeituosa	Correta
	18 (VP)	2 (FN)
	5 (FP)	75 (VN)

Total: 100 peças.

- Entre as peças que o modelo disse que eram defeituosas, quantas realmente eram defeituosas?

$$Precisão = \frac{VP}{VP + FP}$$

$$Precisão = \frac{18}{18 + 5} = \frac{18}{23} = 0,78 = 78\%$$

- Contexto: evita o desperdício de peças.

Recall (sensibilidade)

Real	Previsto	
	Defeituosa	Correta
	Defeituosa	Correta
	18 (VP)	2 (FN)
	5 (FP)	75 (VN)

Total: 100 peças.

- Entre as peças que o modelo disse que eram defeituosas, quantas o modelo acertou na classificação?

$$Recall = \frac{VP}{VP + FN}$$

$$Recall = \frac{18}{18 + 2} = \frac{18}{20} = 0,90 = 90\%$$

- Contexto: evita enviar para o cliente peças com defeito.

Especificidade

Previsto

<i>Real</i>		Defeituosa	Correta
	Defeituosa	18 (VP)	2 (FN)
	Correta	5 (FP)	75 (VN)

Total: 100 peças.

- Entre as peças que realmente estavam boas, quantas o modelo classificou corretamente como boas?

$$Espec = \frac{VN}{VN + FP}$$

$$Espec = \frac{75}{75 + 5} = \frac{75}{80} = 0,94 = 94\%$$

- Contexto: evita o desperdício de peças.

Fall-out ou Taxa de Falsos Positivos

Previsto

<i>Real</i>		Defeituosa	Correta
	Defeituosa	18 (VP)	2 (FN)
	Correta	5 (FP)	75 (VN)

Total: 100 peças.

- Mede a proporção de peças ruins que foram classificadas como boas incorretamente.

$$TFP = \frac{FP}{FP + VN}$$

$$TFP = \frac{5}{5 + 75} = \frac{5}{80} = 0,06 = 6\%$$

- TFP também pode ser calculada como: **1 - Espec.**

Medida F1 (*f*-score)

		Previsto	
Real		Defeituosa	Correta
	Defeituosa	18 (VP)	2 (FN)
	Correta	5 (FP)	75 (VN)

Total: 100 peças.

- Combina Precisão e Recall (sensib.) numa única métrica (média harmônica):

$$F1 = 2 \times \frac{\text{Precisão} \times \text{Recall}}{\text{Precisão} + \text{Recall}}$$

$$F1 = 2 \times \frac{0,78 \times 0,90}{0,78 + 0,90} = 0,84 = 84\%$$

- Contexto: O F1 mostra o equilíbrio entre não deixar defeitos passarem (recall) e não acusar defeito à toa (precisão).

Resumindo...

- Acurácia (93%)**: o modelo acerta bastante no geral.
- Precisão (78%)**: quando acusa defeito, geralmente está certo, mas ainda descarta algumas peças boas.
- Recall (90%)**: impede de enviar peças com defeitos (bom para segurança e qualidade).
- Especificidade (94%)**: reconhece bem as peças boas, evitando desperdício.
- F1 (84%)**: equilíbrio razoável entre encontrar defeitos e não exagerar nos falsos alarmes.

DÚVIDAS?

Curva ROC

- ROC = *Receiver Operating Characteristic* = Característica operacional do receptor;
- É um **gráfico** que mostra o desempenho de um modelo de classificação em diferentes **limiares de decisão**;
- Invés de olhar só para um ponto fixo (matriz de confusão), a curva ROC avalia o comportamento do modelo à medida que variamos o limite de decisão.

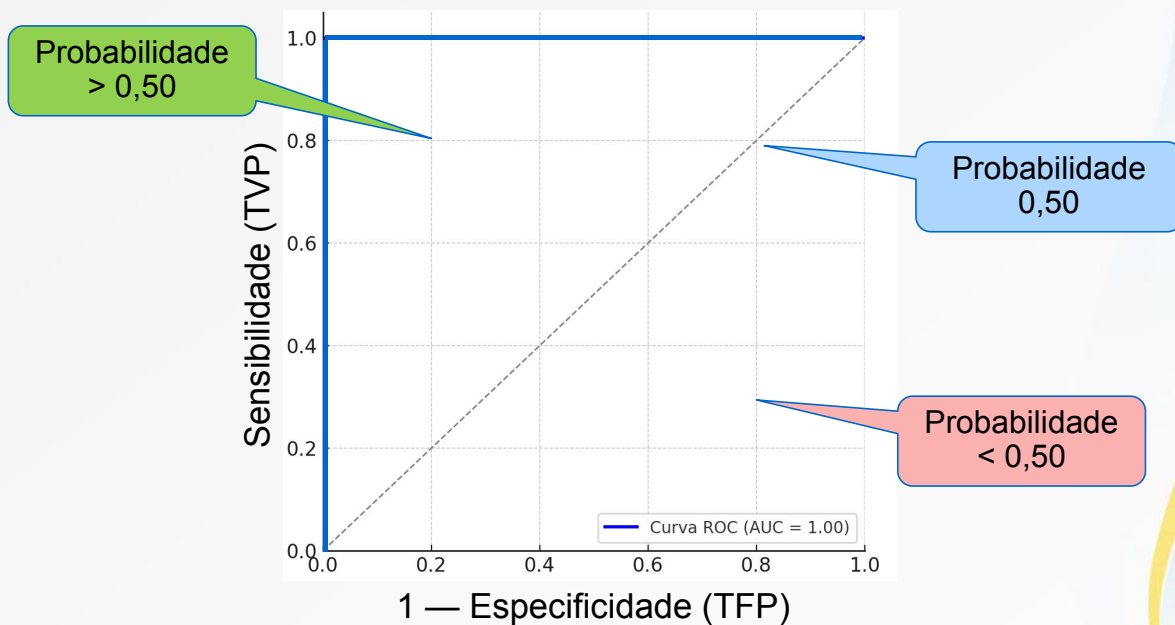
Curva ROC - Etapas de construção

1. O modelo normalmente dá uma probabilidade de que cada peça seja defeituosa.
 - Ex.: “essa peça tem 80% de chance de estar com defeito”.
2. Para decidir se a peça é Defeituosa (1) ou Correta (0), escolhemos um limiar (*threshold*).
 - Se probabilidade \geq limiar: classifica como “Defeituosa”.
 - Se probabilidade $<$ limiar: classifica como “Correta”.

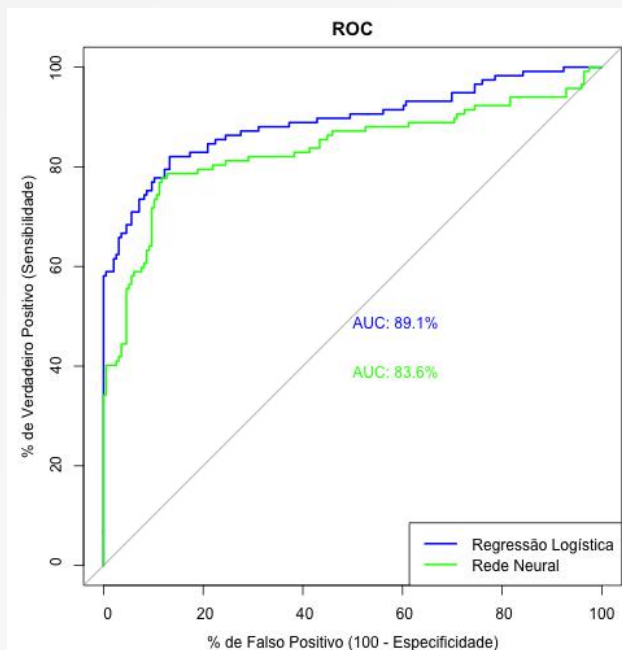
Curva ROC - Etapas de construção

3. Para cada limiar (0,0 até 1,0), calculamos:
 1. Taxa de Verdadeiros Positivos (*recall* ou sensibilidade);
 2. Taxa de Falsos Positivos (*fall-out* ou 1 - especificidade).
4. Cada par (TFP, TVP) é um ponto no gráfico;
5. A curva ROC é a linha que liga esses pontos.

Curva ROC - Como interpretar?



Curva ROC do mundo real



- AUC: *area under curve*;
- O ideal é AUC = 1 (100%), mas, entre 0,75 e 0,95 está muito aceitável.

Universidade Federal do Pará
Programa de Pós-graduação em Engenharia de Processos
(PPGEP)
Disciplina: Tópicos Especiais em Engenharia de Processos
Prof. Dr. Alan de Souza
Aula 4

Setembro/2025

Métricas de avaliação de modelos de regressão

1. SSE (Sum of Squared Error / Soma dos Erros Quadráticos)
2. MAE (Mean Absolute Error / Erro Absoluto Médio)
3. MSE (Mean Squared Error / Erro Quadrático Médio)
4. RMSE (Root Mean Squared Error / Raiz do Erro Quadrático Médio)
5. MAPE (Mean Absolute Percentage Error / Erro Percentual Absoluto Médio)
6. SMAPE (Symmetric MAPE)

SSE (Sum of Squared Error / Soma dos Erros Quadráticos)

- Também chamada de RSS (Residual Sum of Squares)
- Mede o erro total absoluto cometido pelo modelo (sem normalização).

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

MAE (Mean Absolute Error / Erro Absoluto Médio)

- Fácil de interpretar;
- Não penaliza fortemente grandes erros.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

MSE (Mean Squared Error / Erro Quadrático Médio)

- Penaliza fortemente grandes erros (quadrado);
- Muito utilizado na literatura.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

RMSE (Root Mean Squared Error / Raiz do Erro Quadrático Médio)

- Métrica na mesma unidade da variável dependente;
- Mais interpretável que o MSE.

$$RMSE = \sqrt{MSE}$$

MAPE (Mean Absolute Percentage Error / Erro Percentual Absoluto Médio)

- Erro médio em porcentagem.
- Problema: gera valores muito altos quando a saída real (y_i) é muito próxima de zero e quando n é pequeno.

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

SMAPE (Symmetric MAPE)

- Versão simétrica do MAPE que evita distorções quando y_i é muito pequeno.

$$SMAPE = \frac{100}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)/2}$$

DÚVIDAS?

Projetos mais robustos

- Projeto sobre regressão de nota de admissão em uma universidade americana.
- Projeto sobre classificação de fluidos Newtonianos e não-Newtonianos.
- Link dos projetos: <https://github.com/amarcel/ML-PPGEP-set2025/>