



ESCOLA POLITÉCNICA DA UNIVERSIDADE DE SÃO PAULO

PQI 3403 Análise de Processos da Industria Química

Ardson dos Santos Vianna Júnior - ASVJ
e-mail: ardson@usp.br





ESCOLA POLITÉCNICA DA UNIVERSIDADE DE SÃO PAULO

Aula Machine learning

PQI 3403 Análise de Processos da Indústria Química



Sumário

- Introdução
- Matriz de confusão
- ROC
- Conclusão

Matriz de confusão

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

Exemplo – a imagem é um gato?

- Matriz de confusão

Real / predito	Gato	Não gato
gato	TP = 25	FP=10
não gato	FN = 25	TN = 40

Desempenho

- Acurácia: $= \frac{TP+TN}{TP+TN+FN+FP} = \frac{TP+TN}{TP+TN+FN+FP} = \frac{25+40}{100} = 0,65$

- Precisão: $= \frac{TP}{TP+FP} = \frac{25}{25+25} = 0,5$

- Especificidade: $= \frac{TN}{TN+FP} = \frac{40}{40+10} = 0,8$

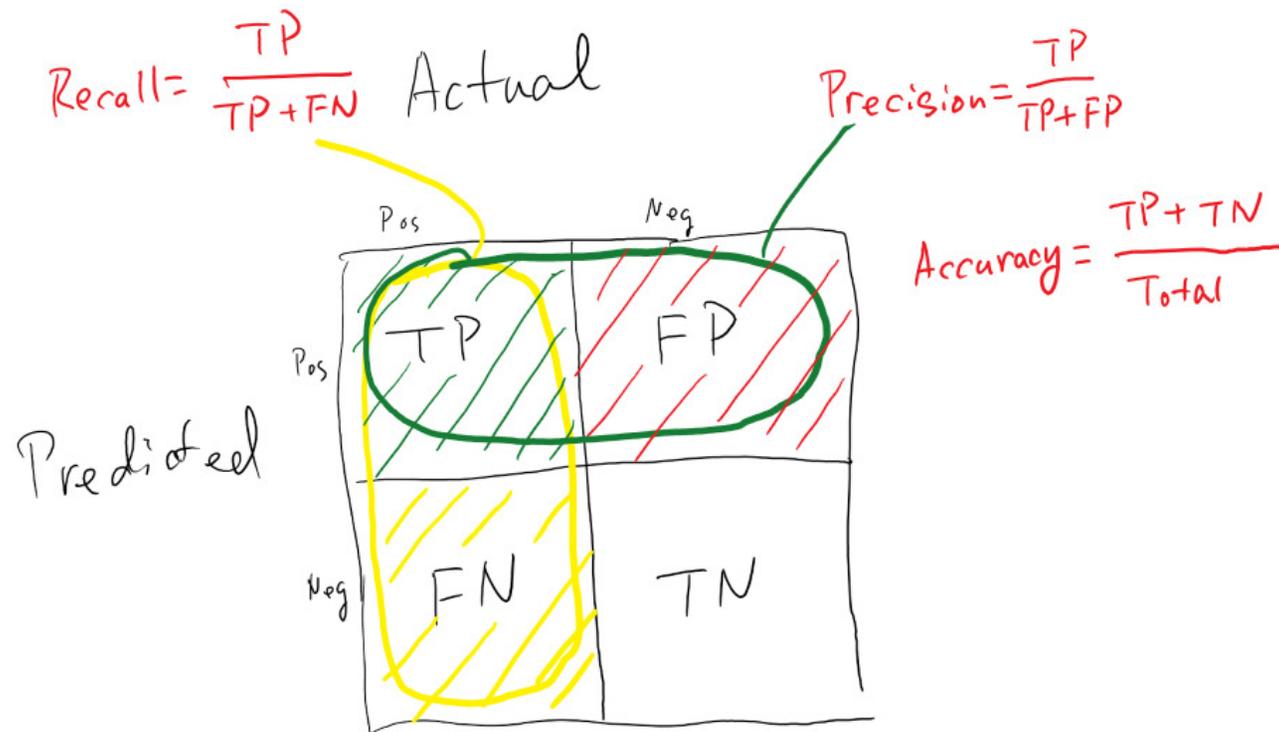
- Recall (sensibilidade): $= \frac{TP}{TP+FN} = \frac{25}{25+25} = 0,5$

-

Real / predito	Gato	Não gato
gato	TP = 25	FP=10
não gato	FN = 25	TN = 40

parâmetros

<https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62>



Classificação - desempenho

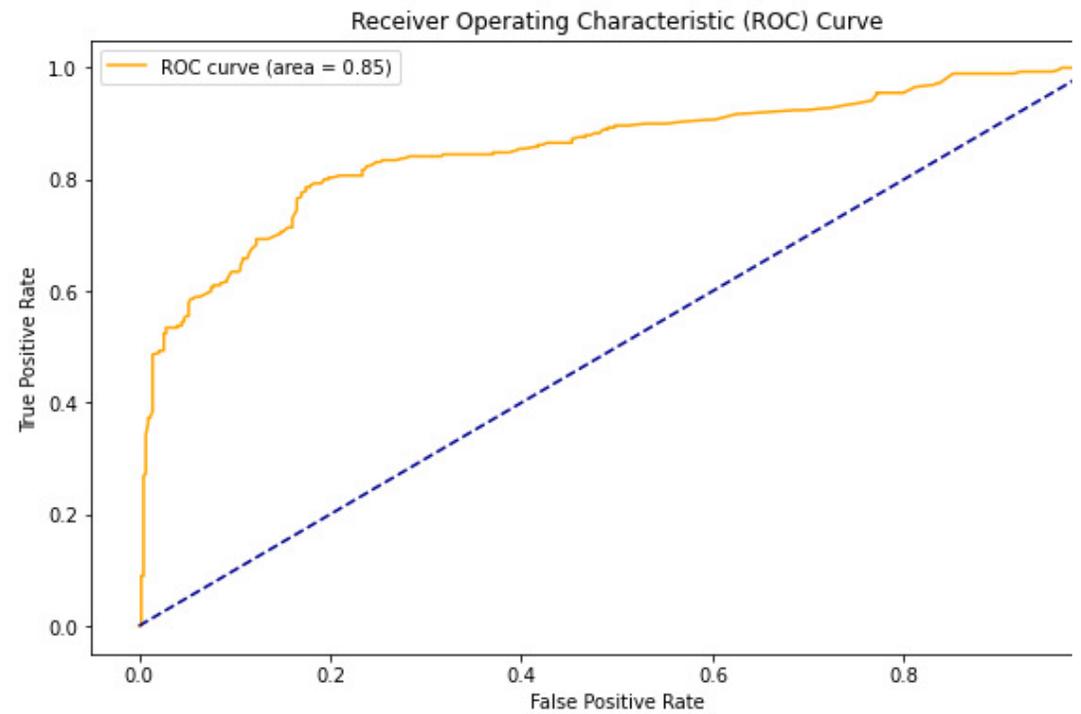
- A Curva Característica de Operação do Receptor (Curva COR), ou, do inglês, **Receiver Operating Characteristic Curve (ROC curve)**
- TPR (true positive rate) e FPR (false positive rate)
- $TPR = TP / (TP + FN)$
- $FPR = FP / (FP + TN)$

Classificação - desempenho

- A Curva Característica de Operação do Receptor (Curva COR) - (**ROC curve**)
- $TPR = TP / (TP + FN)$
- $FPR = FP / (FP + TN)$
- Compromisso entre verdadeiros positivos e falsos negativos
- Não é sensível ao desbalanço de classes

ROC

- Probabilidades
- Limiar da classificação



Classificação - desempenho

- AUC – *area under ROC curve* área sobre a curva ROV
- É uma medida acumulada da performance de todas as possibilidades de limitação.



Conclusão

- Matriz de confusão
- Parâmetros de desempenho
- Curva ROC

Bibliografia

- <https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-867-machine-learning-fall-2006/index.htm>
- Guttag, John. *Introduction to Computation and Programming Using Python: With Application to Understanding Data*. 2nd ed. MIT Press, 2016. ISBN: 9780262529624
- <https://www.youtube.com/watch?v=h0e2HAPTGF4>