

Aprendizado de Máquina: Medidas para Avaliação de Classificadores

Prof. Arnaldo Candido Junior
UTFPR – Medianeira

Matriz de confusão

- Acurácia trata as classes igualmente, pode não ser adequada para dados desbalanceados
- Em muitos problemas, classe rara é mais interessante que a majoritária.
 - Ex.: diagnóstico de doenças
- Veremos medidas alternativas baseadas na matriz de confusão, muitas das quais analisadas em pares

Matriz de confusão ₍₂₎

- Matriz de confusão: classe real nas colunas e predições nas linhas (convenção)

Classe Predita	Classe Verdadeira	
	Positiva	Negativa
Positiva	Verdadeiro Positivo	Falso Positivo
Negativa	Falso Negativo	Verdadeiro Negativo



VP	FP
FN	VN

Matriz de confusão ⁽³⁾

- Verdadeiro positivo: instância da classe positiva prevista como classe positiva
- Falso positivo: instância que não é da classe positiva prevista como classe positiva
- Falso negativo: instância da classe positiva que não foi prevista como positiva
- Verdadeiro negativo: instância que não é da classe positiva e que não foi prevista como positiva

Matriz de confusão ₍₄₎

- Classe positiva é, em geral, a classe de maior interesse (ou classe com menos exemplos)
 - Em alguns casos, os erros têm igual importância
 - Em outro casos, erros diferentes têm consequências diferentes
 - Por exemplo, falso negativo é pior no diagnóstico de doenças

Medidas

- Acurácia ou taxa de acerto

$$ACC = \frac{VP + VN}{VP + FP + FN + VN}$$

VP	FP
FN	VN

- Convenção: 

Medidas ₍₂₎

- Taxa de erro

$$ERR = \frac{FP + FN}{VP + FP + FN + VN}$$

VP	FP
FN	VN

$$ERR = 1 - ACC$$

Medidas ₍₃₎

- Taxa de falsos positivos (em relação aos negativos)

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

VP	FP
FN	VN

- Também chamada de: Erros do Tipo 1, Custo

Medidas ₍₄₎

- Taxa de falsos negativos (em relação aos positivos)

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

VP	FP
FN	VN

- Também chamada de: Erros do Tipo 2

Medidas ₍₅₎

- Taxa de verdadeiros positivos (em relação aos positivos)

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

VP	FP
FN	VN

- Também chamada de: Benefício, Sensibilidade, Cobertura, Revocação, Recall

Medidas ₍₆₎

- Precisão (em relação às predições positivas)

$$P = \frac{VP}{VP + FP}$$

VP	FP
FN	VN

- Especificidade (em relação às predições negativas)

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

VP	FP
FN	VN

Medidas ₍₇₎

- Medida-F: média hârmônica ponderada entre a precisão e a cobertura

$$F = \frac{(1 + \alpha) P C}{\alpha P + C}$$

- Onde α representa o peso dado a precisão

Medidas ₍₈₎

- Medida- F_1 : variante da Medida-F em que precisão e cobertura tem o mesmo peso

$$F_1 = \frac{2PC}{P+C} = \frac{2}{\frac{1}{P} + \frac{1}{C}}$$

- Define uma média harmônica entre precisão e cobertura
- Obriga classificador a equilibrar precisão e cobertura ou terá Medida-F baixa

Análises

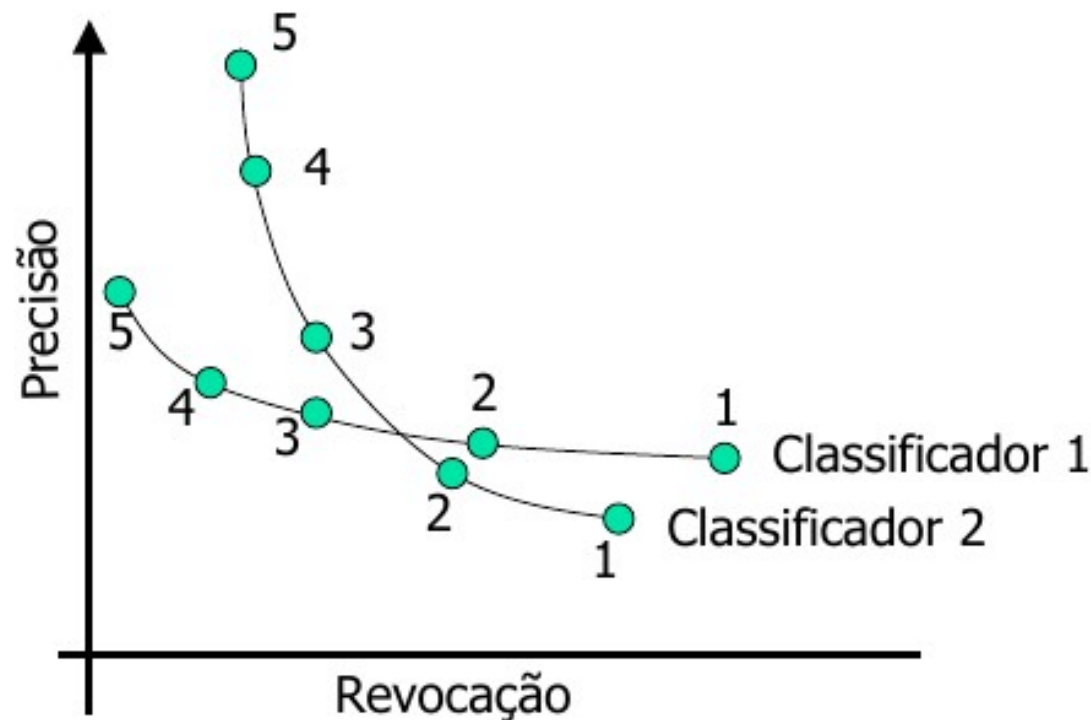
- As medidas são usadas popularmente em diferentes áreas
 - Recuperação de Informações: Precisão vs Cobertura (**foco**)
 - Estatística: Erros do Tipo 1 vs Erros do Tipo 2
 - Biologia, Medicina: Especificidade vs Sensibilidade
 - Medicina, Processamento de Sinais: TVP vs TFP (ou Custo vs Benefício) (**foco**)

Cobertura vs Precisão

- **Cobertura** (*recall* ou revocação): **tudo que é relevante foi recuperado?** Nenhum exemplo positivo é deixado de fora
- **Precisão**: **tudo que foi recuperado é relevante?** Nenhum exemplo negativo é incluído
- Em recuperação de informações, VN costuma ser grande e atrapalhar o cálculo de diversas medidas
 - Precisão e Cobertura não usam VN

Cobertura vs Precisão ₍₂₎

- É possível ajustar um classificador para focar mais em precisão ou cobertura mudando o limiar do score do classificador para decidir entre as classes



Exercícios

- É possível calcular as medidas classe a classe gerando várias matrizes de confusão
- Exercício 1: gerar matriz de confusão chamando uma das classes positiva e as outras duas de negativas
- Exercício 2: idem para as outras duas classes

Classe predita	Classe verdadeira		
	1	2	3
1	25	10	0
2	0	40	0
3	5	0	20

Exercícios ₍₂₎

- Exercício 3: calcular medidas para:

		Classe verdadeira	
		P	N
Classe predita	p	25	10
	n	45	60

Classificador 1

TVP =

TFP =

		Classe verdadeira	
		P	N
Classe predita	p	70	20
	n	15	30

Classificador2

TVP =

TFP =

		Classe verdadeira	
		P	N
Classe predita	p	70	95
	n	30	5

Classificador 3

TVP =

TFP =

Exercícios ₍₃₎

- Exercício 4: seja um classificador com a seguinte matriz de confusão, definir:
 - Acurácia, Precisão, Cobertura (sensibilidade) e Especificidade

		Classe verdadeira	
		P	N
Classe predita	p	70	40
	n	30	60

Gráficos ROC

- Do inglês, Receiver Operating Characteristics
- Medida de desempenho originária da área de processamento de sinais
- Plotar no gráfico ROC os 3 classificadores do exemplo anterior

Classificador 1
TVP = 0.4
TFP = 0.3



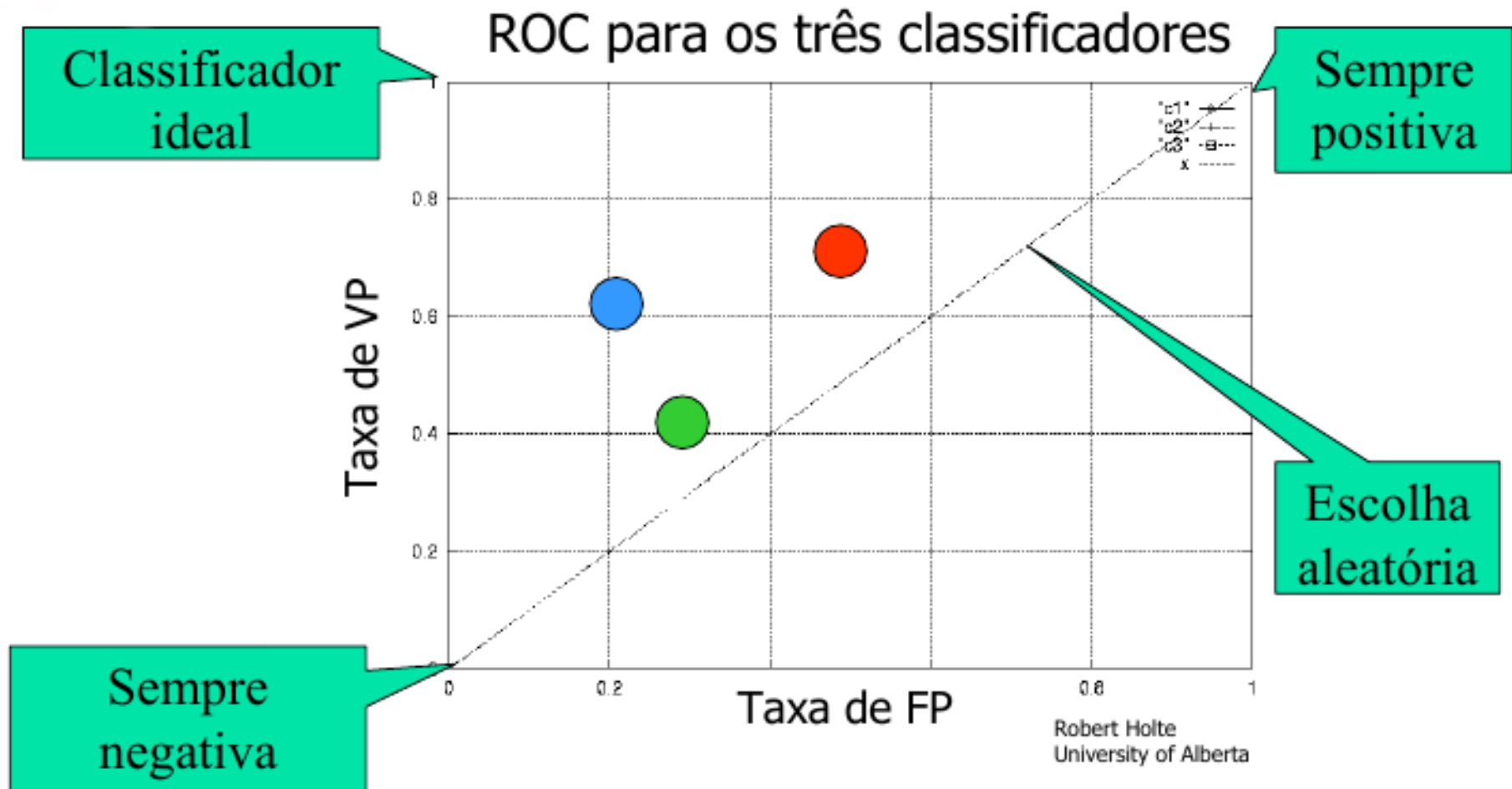
Classificador2
TVP = 0.7
TFP = 0.5



Classificador 3
TVP = 0.6
TFP = 0.2



Gráficos ROC



Gráficos ROC ₍₂₎

- Informalmente, melhor classificador é aquele cujo ponto está mais a noroeste
- Classificadores próximos do canto inferior esquerdo são conservadores. Só fazem classificações positivas com forte evidência (poucos erros de FP)
- Classificadores próximos ao canto superior direito são liberais (correm riscos) (muitos FPs)

Gráficos ROC ₍₃₎

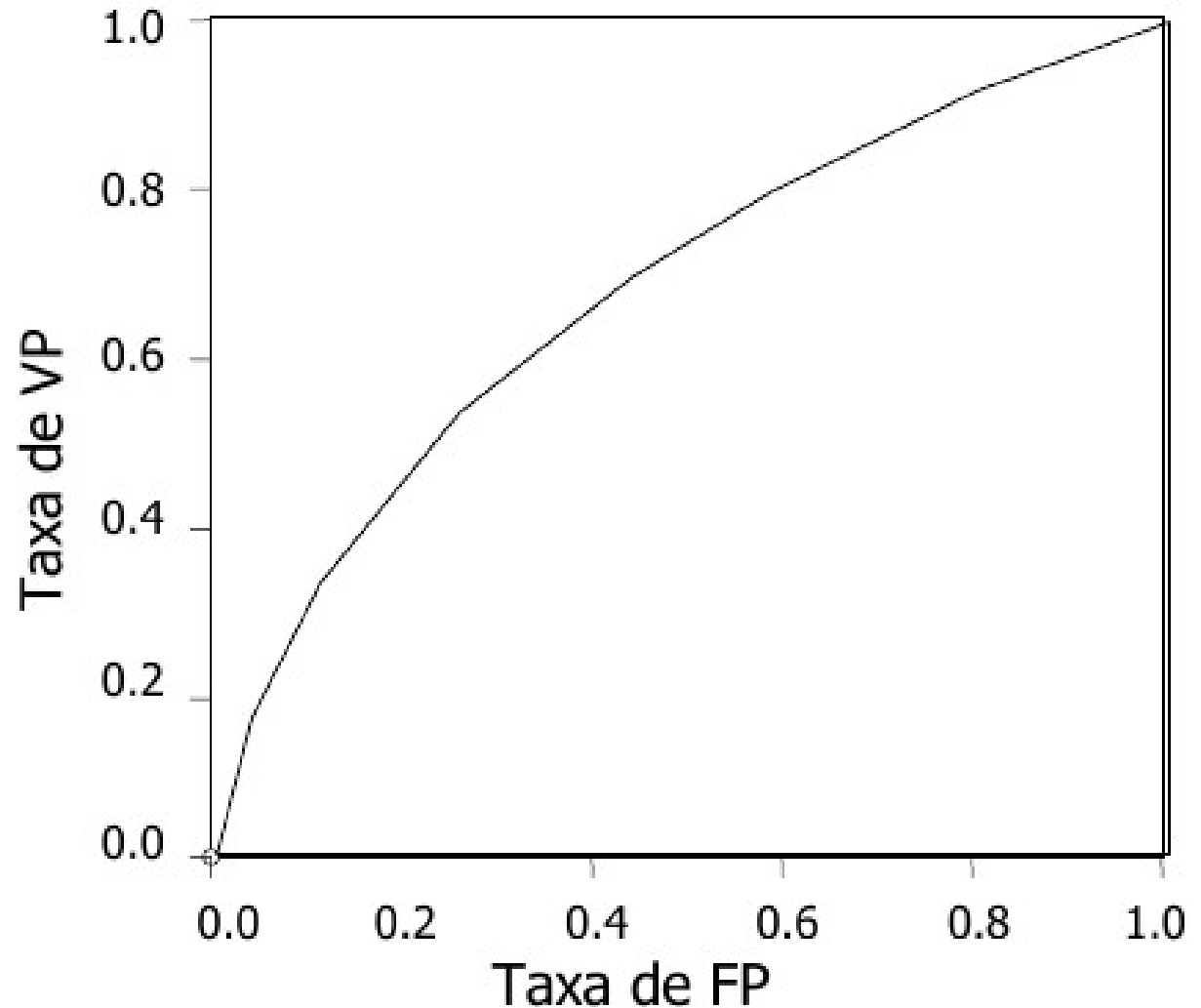
- Pontos na diagonal estão associados à escolha aleatória da classe
 - Instância é classificada como positiva com probabilidade fixa
 - Independente dos valores de seus atributos de entrada
 - TVP e TFP serão idênticas

Gráficos ROC ₍₅₎

- Classificadores que geram escores ou probabilidades
 - Diferentes valores de limiares podem ser utilizados escolha da classe positiva, gerando vários pontos
 - Cada valor de limiares produz um ponto diferente
 - Ligação dos pontos gera uma curva ROC

Gráficos ROC ⁽⁶⁾

Classificador Escore/
Probabilístico



Curvas ROC ₍₂₎

1 - Ordenar exemplos em ordem decrescente por valor de predição

2- Aplicar threshold a cada valor de predição

3 – Calcular VP, VN, FP, FN para cada threshold

4 – Calcular TVP e TFP

$$Classe = \begin{cases} \text{predição} \geq \theta, P \\ \text{predição} < \theta, N \end{cases}$$

Instância	Clas-Ver	Predição
6	P	0.9
3	P	0.8
2	N	0.7
9	P	0.6
5	P	0.6
1	N	0.5
7	N	0.3
8	N	0.2
4	N	0.2
10	N	0.1

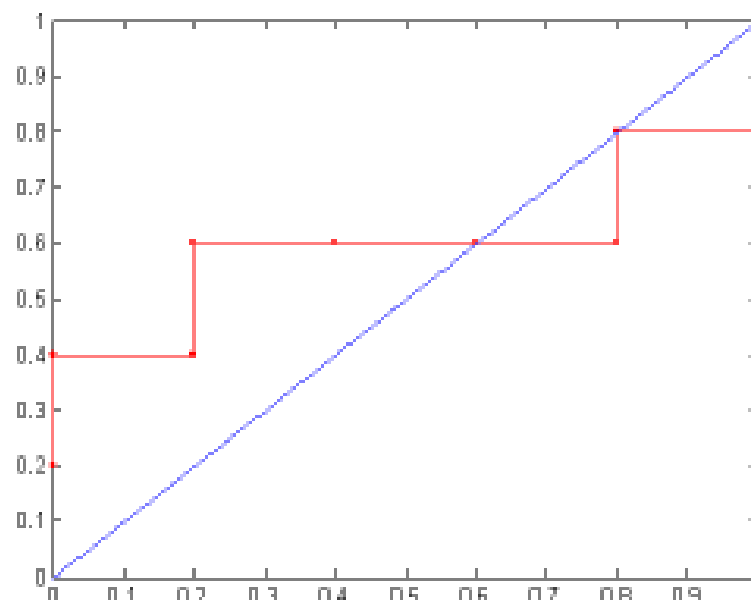
Curvas ROC ₍₃₎

Threshold \geq

Classe	+	-	+	-	-	-	+	-	+	+	
P	0.25	0.43	0.53	0.76	0.85	0.85	0.85	0.87	0.93	0.95	1.00
VP	5	4	4	3	3	3	3	2	2	1	0
FP	5	5	4	4	3	2	1	1	0	0	0
VN	0	0	1	1	2	3	4	4	5	5	5
FN	0	1	1	2	2	2	2	3	3	4	5
TVP	1	0.8	0.8	0.6	0.6	0.6	0.6	0.4	0.4	0.2	0
TFP	1	1	0.8	0.8	0.6	0.4	0.2	0.2	0	0	0

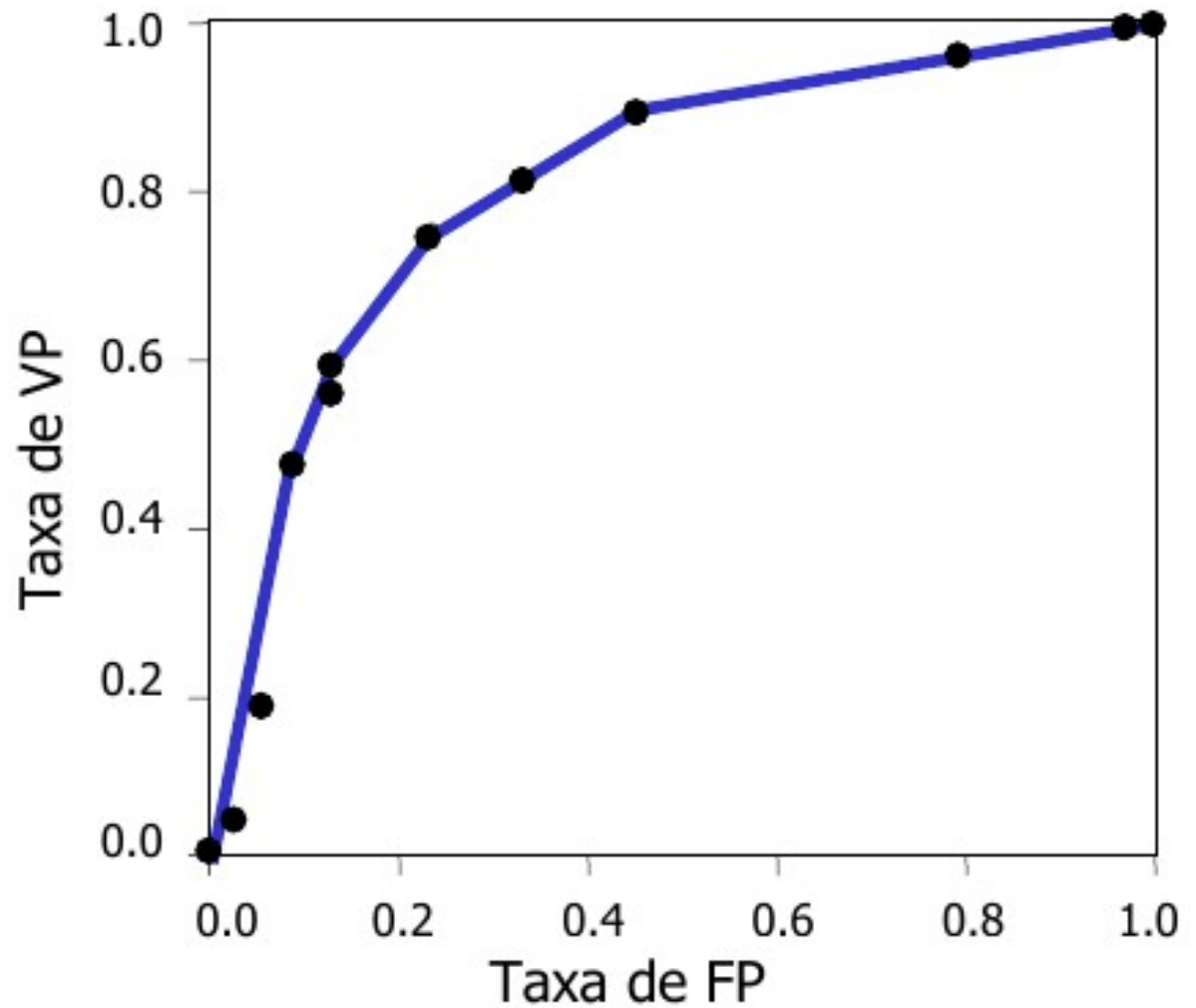


Curva ROC:



Curvas ROC ₍₄₎

Classificador
Discreto

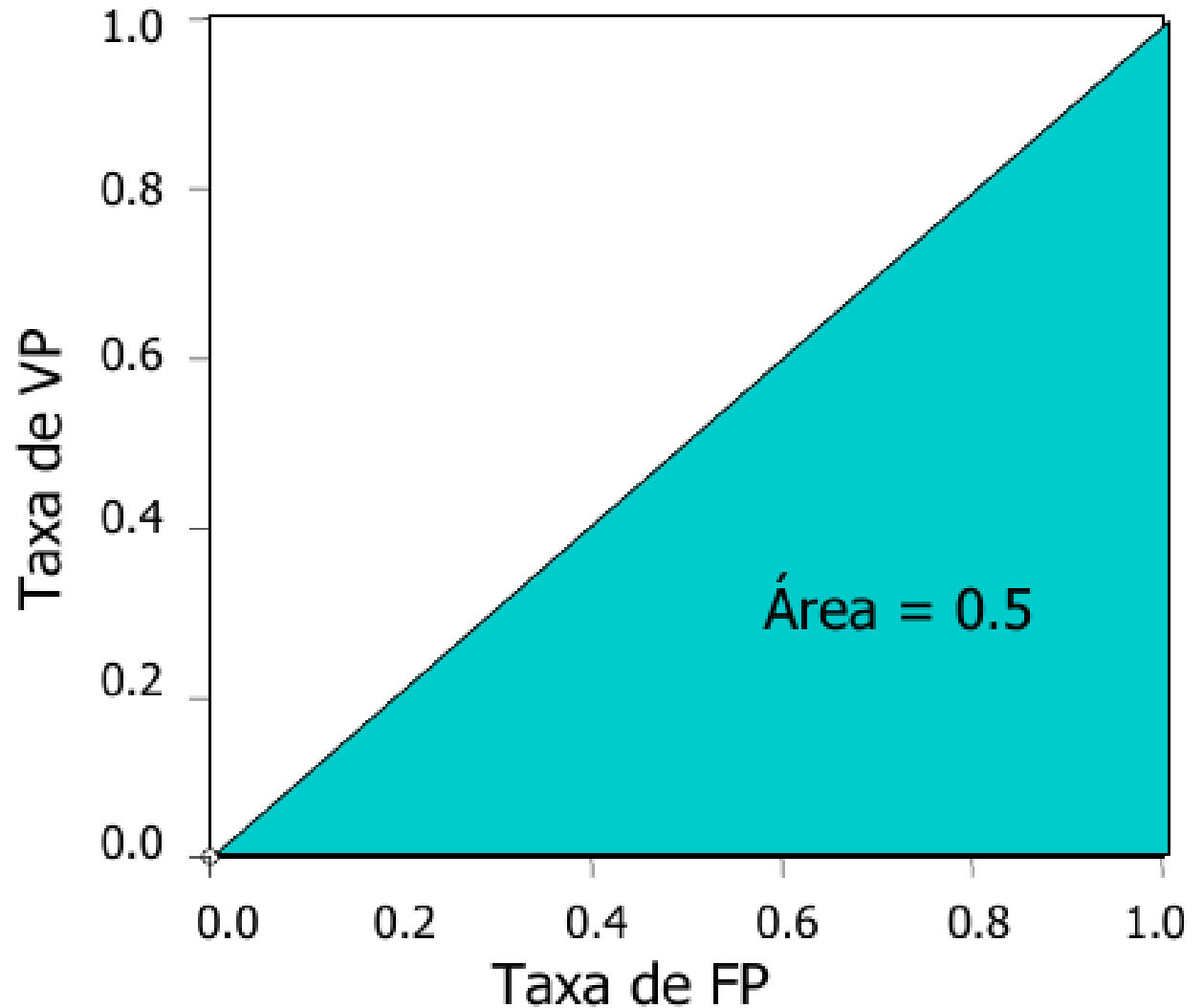


Área sob a Curva ROC (AUC)

- Fornece uma estimativa do desempenho de classificadores
 - Gera um valor contínuo no intervalo $[0, 1]$
 - Quanto maior melhor
 - Cálculo: somar as áreas de sucessivos trapezoides
- Um classificador com maior AUC pode apresentar AUC pior em trechos da curva
- É mais confiável utilizar médias de curvas

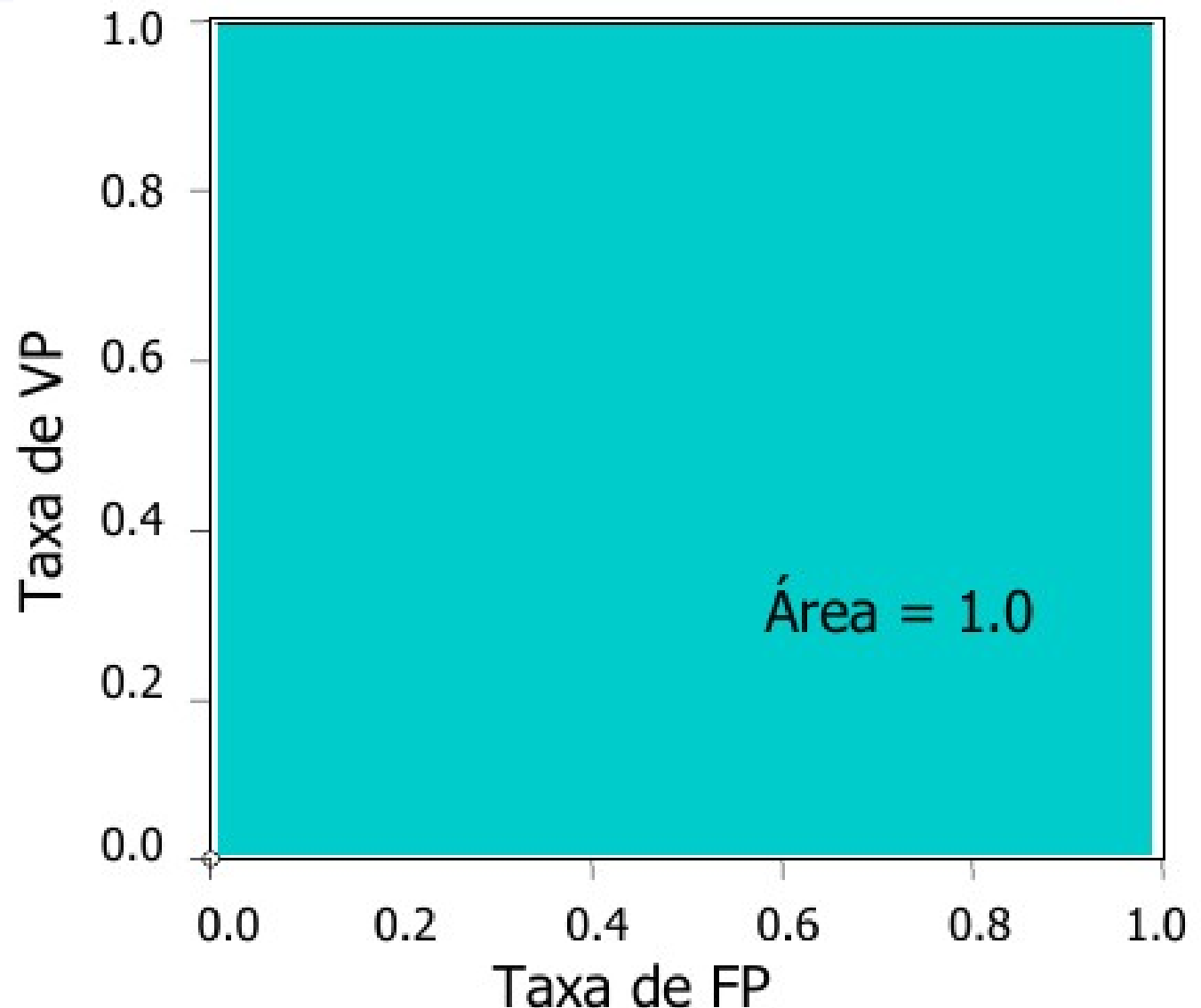
Área Sob Curvas ROC ⁽²⁾

Nenhuma
Discriminação

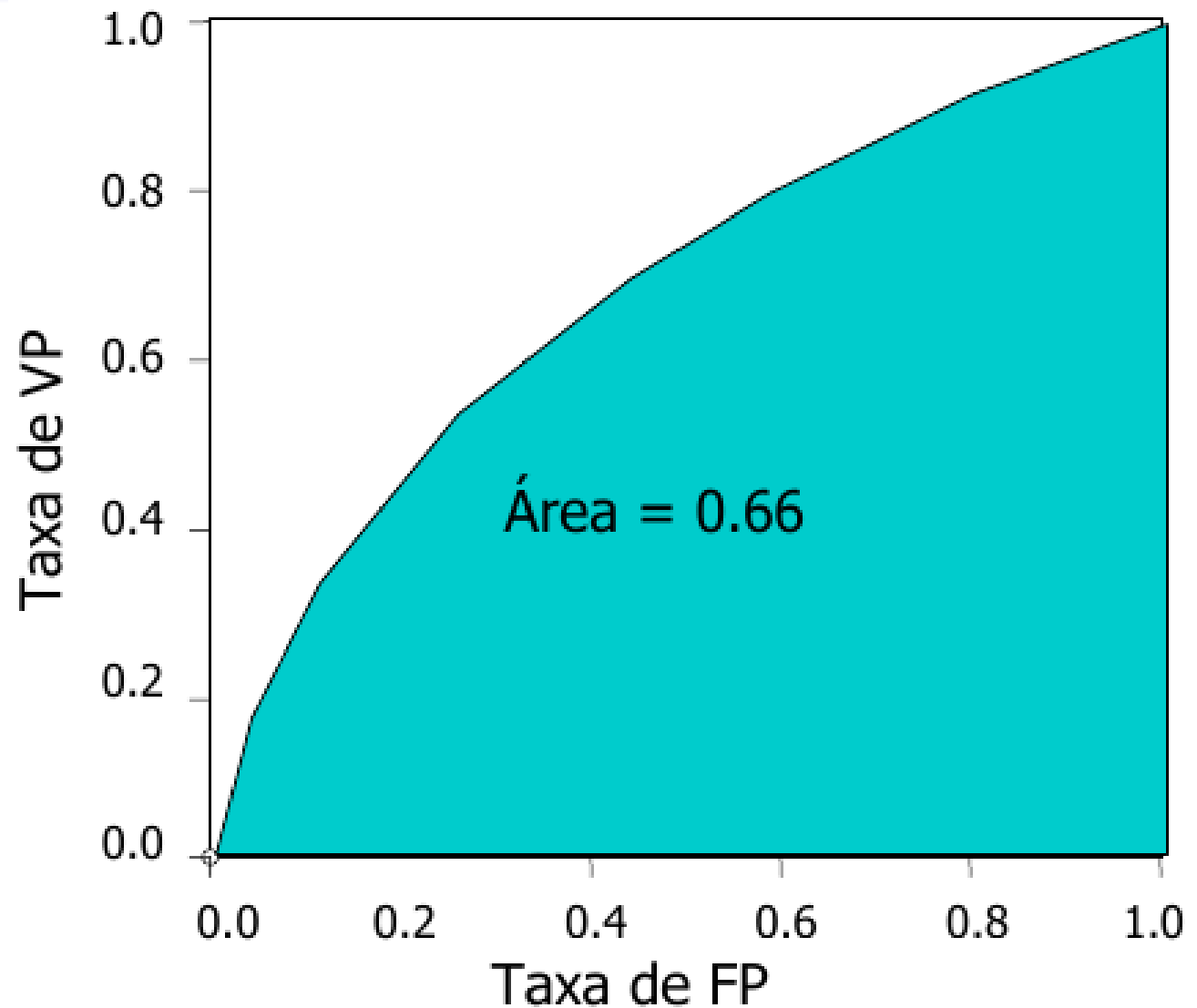


Área Sob Curvas ROC ⁽³⁾

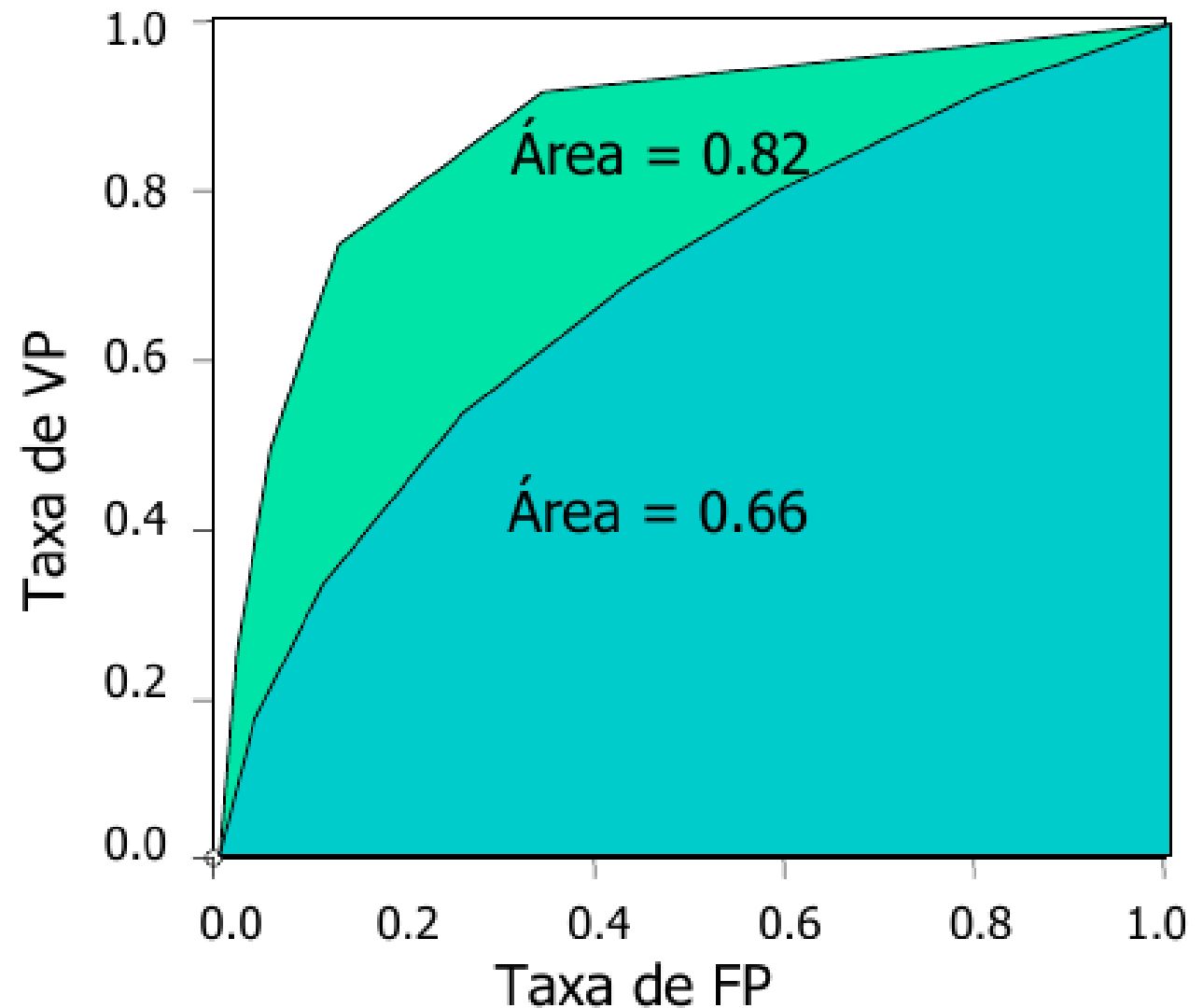
Discriminação
Perfeita



Área Sob Curvas ROC ⁽⁴⁾



Área Sob Curvas ROC ⁽⁵⁾



Créditos

- Adaptado de:
 - Notas de aula do Prof. Dr. André C. P. L. F. de Carvalho - ICMC-USP