



Departamento de Engenharia Elétrica

SEL-0339

Introdução à Visão Computacional

Aula 11

Avaliação de Classificadores

Prof. Dr. Marcelo Andrade da Costa Vieira

Arthur Chaves Costa

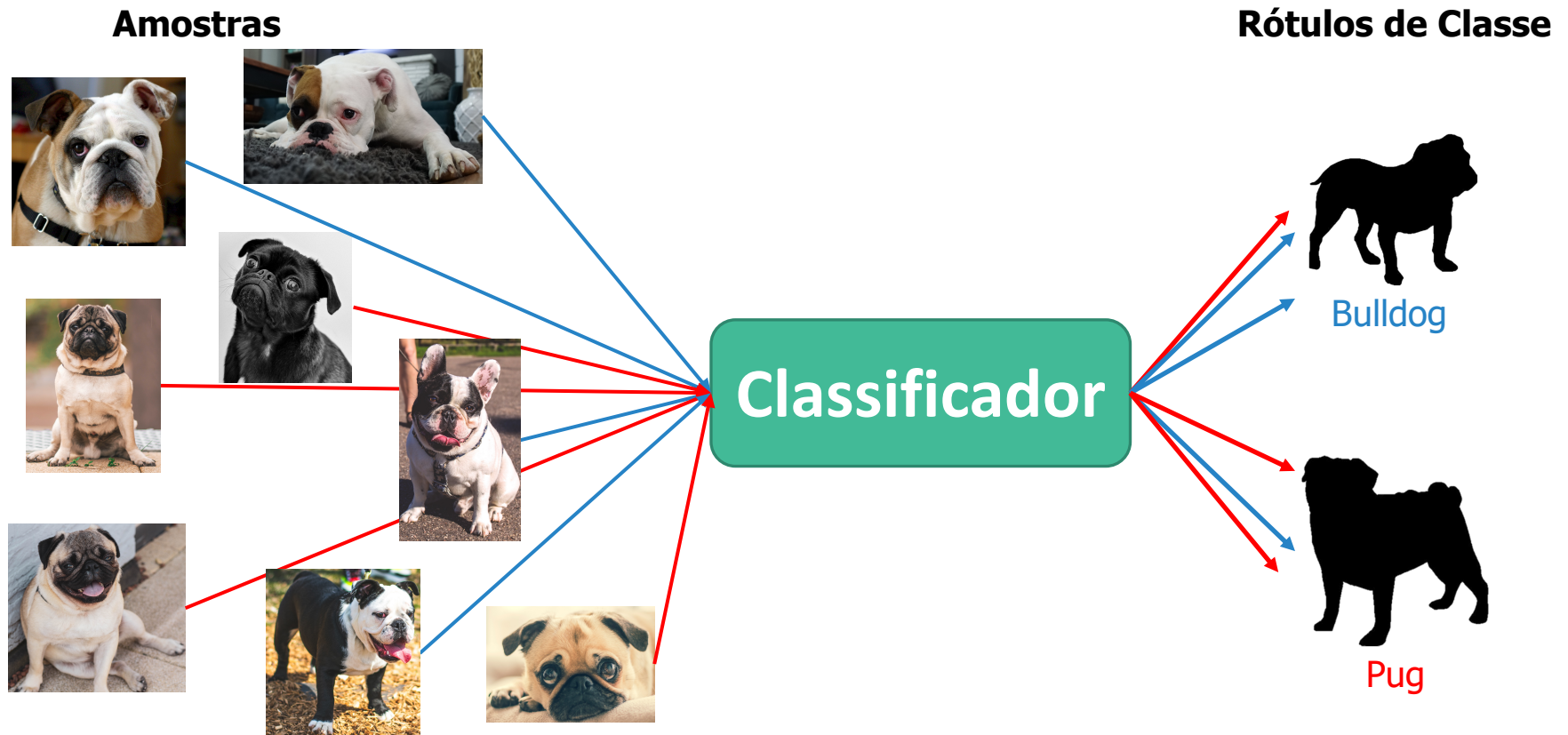
Renann de Faria Brandão

mvieira@sc.usp.br

Função de um Classificador:

Mapeamento → Atribuir uma **Instância** a uma **Classe**.

Cada amostra é mapeada para um elemento do conjunto de rótulos de classe como positiva (p) ou negativa (n).



Tipos de Classificadores:

Classificador Discreto → Atribui a classe diretamente.

Classificador Contínuo → Atribui uma “probabilidade” do elemento ser membro de uma classe. Um “Threshold” determina a classe.



Validação Cruzada

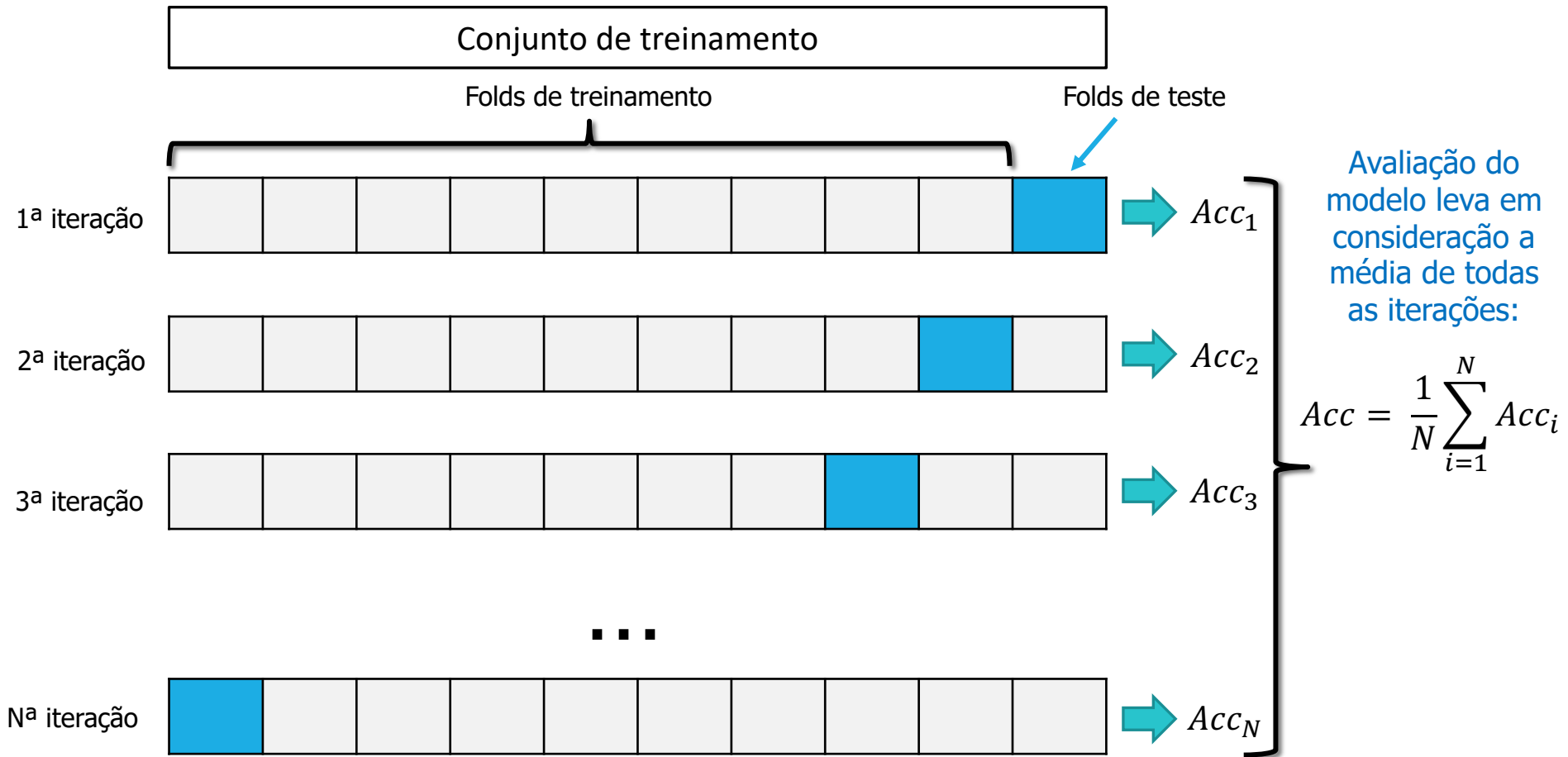
AVALIAÇÃO DE CLASSIFICADORES

Validação Cruzada

- Métodos utilizados para avaliar de maneira mais robusta os classificadores;
- Avaliam a capacidade de generalização do modelo a partir de um conjunto de dados;
- Principalmente utilizadas em casos com quantidade de dados limitada;
- O conceito central é o particionamento do conjunto de dados em vários subconjuntos que são utilizados de maneira revezada na estimação das métricas;
- Os métodos mais utilizados são:
 - K-fold
 - Subamostragem aleatória repetitiva (*holdout*)

K-Fold

Revezamento de subconjuntos mutualmente excludentes.



Subamostragem aleatória repetitiva (*holdout*)

Realiza-se a divisão dos conjuntos de treinamento e teste (ex. 70% e 30%) de maneira aleatória por N vezes, e em seguida realiza-se a média das métricas de avaliação.

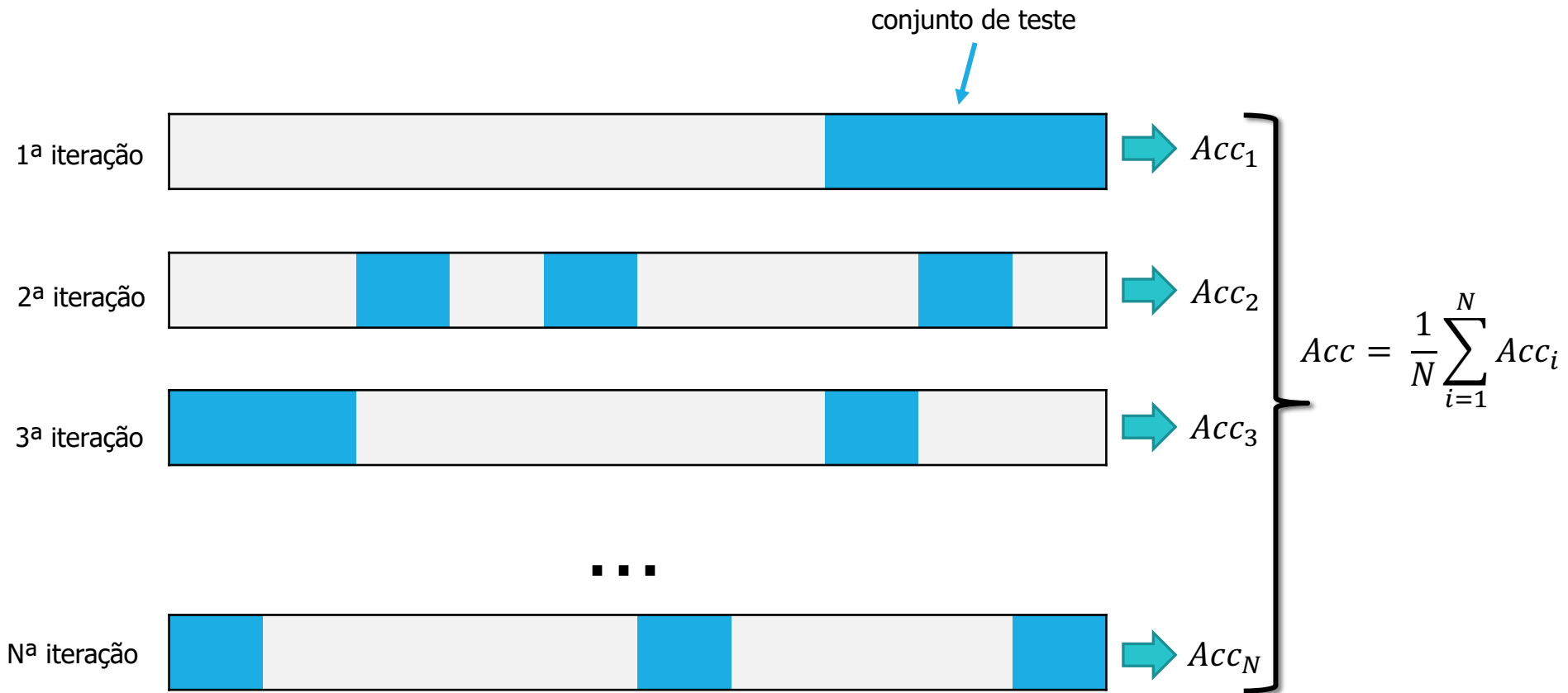


Tabela de Contingências.

Saída\Entrada	Genuíno (Positivo)	Impostores (Negativos)	
Positivo	VP(TP)	FP	TSP(B)
Negativo	FN	VN(TN)	TSN
	TG(A)	TI	TS

- Saída: **Positivo** → O classificador classifica a amostra como pertencente à classe.
- Saída: **Negativo** → O classificador classifica a amostra como não pertencente à classe.

Tabela de Contingências.

Saída\Entrada	Genuíno (Positivo)	Impostores (Negativos)	
Positivo	VP(TP)	FP	TSP(B)
Negativo	FN	VN(TN)	TSN
	TG(A)	TI	TS

- VP(TP) → VERDADEIRO POSITIVO (True Positive, Hit) → O elemento de entrada é Genuíno (Positivo) e o Classificador o classifica como Positivo.
- VN (TN) → VERDADEIRO NEGATIVO (True Negative, Correct Rejection) → O elemento de entrada é Impostor (Negativo) e o Classificador o classifica como Negativo.
- FP → FALSO POSITIVO (False Positive, False Alarm, Type I error) → O elemento de entrada é Impostor (Negativo) e o Classificador o classifica como Positivo.
- FN → FALSO NEGATIVO (False Negative, Type II error) → O elemento de entrada é Genuíno (Positivo) e o classificador o classifica como Negativo.

Tabela de Contingências.

Saída\Entrada	Genuíno (Positivo)	Impostores (Negativos)	
Positivo	VP(TP)	FP	TSP(B)
Negativo	FN	VN(TN)	TSN
	TG(A)	TI	TS

- TG → Total de Genuínos (Total de elementos de uma classe) → $TG = TP + FN$
- TI → Total de Impostores → $TI = TN + FP$
- TSP → Total de Saídas Positivas → $TSP = TP + FP$
- TSN → Total de Saídas Negativas → $TSN = TN + FN$
- TCC → Total de Classificações Corretas → $TCC = TP + TN$
- TCI → Total de Classificações Incorretas → $TCI = FP + FN$
- TS → Total de Saídas → $TS = TG + TI = TSP + TSN = TP + FP + TN + FN$

Características de um Classificador (Métricas):

SENSIBILIDADE:

- Proporção de Verdadeiros Positivos classificados corretamente como Genuínos (True Positive Rate -TPR, Hit Rate, Recall).
- Proporção de Genuínos de uma classe que foram classificados como Genuínos.

$$TPR = R = \frac{TP}{TG} = \frac{TP}{TP + FN} \qquad 0 \leq TPR \leq 1$$

Saída\Entrada	Genuíno (Positivo)	Impostores (Negativos)	
Positivo	VP(TP)	FP	TSP(B)
Negativo	FN	VN(TN)	TSN
	TG(A)	TI	TS

Características de um Classificador (Métricas):

ESPECIFICIDADE:

- Proporção de Verdadeiros Negativos classificados corretamente como Impostores (True Negative Rate - TNR)
- Proporção de Impostores que foram classificados como Impostores.

$$TNR = E = \frac{TN}{TI} = \frac{TN}{TN + FP} = 1 - FPR \quad 0 \leq E \leq 1$$

Saída\Entrada	Genuíno (Positivo)	Impostores (Negativos)	
Positivo	VP(TP)	FP	TSP(B)
Negativo	FN	VN(TN)	TSN
	TG(A)	TI	TS

Características de um Classificador (Métricas):

RAZÃO DE FALSOS POSITIVOS:

- Proporção de Falsos Positivos classificados relativamente ao Total de Impostores Existentes (False Positive Rate - FPR, False Alarm Rate, Fall-out).
- Proporção de Impostores erroneamente classificados como Genuínos.

$$FPR = \frac{FP}{TI} = \frac{FP}{TN + FP} \qquad 0 \leq FPR \leq 1$$

Saída\Entrada	Genuíno (Positivo)	Impostores (Negativos)	
Positivo	VP(TP)	FP	TSP(B)
Negativo	FN	VN(TN)	TSN
	TG(A)	TI	TS

Características de um Classificador (Métricas):

RAZÃO DE FALSOS NEGATIVOS:

- Proporção de Falsos Negativos classificados relativamente ao Total de Genuínos Existentes (False Negative Rate - FNR).
- Proporção de amostras Genuínas erroneamente classificadas como Impostoras.

$$FNR = \frac{FN}{TG} = \frac{FN}{TP + FN} \qquad 0 \leq FNR \leq 1$$

Saída\Entrada	Genuíno (Positivo)	Impostores (Negativos)	
Positivo	VP(TP)	FP	TSP(B)
Negativo	FN	VN(TN)	TSN
	TG(A)	TI	TS

Características de um Classificador (Métricas):

VALOR PREVISTO DA SAÍDA POSITIVA (PRECISÃO):

- Proporção de Verdadeiros Positivos classificados.

$$P = \frac{TP}{TSP} = \frac{TP}{TP + FP} \quad 0 \leq P \leq 1$$

VALOR PREVISTO DA SAÍDA NEGATIVA:

- Proporção de Verdadeiros Negativos classificados.

$$N = \frac{TN}{TSN} = \frac{TN}{TN + FN} \quad 0 \leq N \leq 1$$

Saída\Entrada	Genuíno (Positivo)	Impostores (Negativos)	
Positivo	VP(TP)	FP	TSP(B)
Negativo	FN	VN(TN)	TSN
	TG(A)	TI	TS

Características de um Classificador (Métricas):

ACURÁCIA (EXATIDÃO):

- Proporção de Classificações Corretas para o total de elementos classificados (Genuínos e Impostores).

$$A = \frac{TCC}{TS} = \frac{TP + TN}{TS} = \frac{TP + TN}{TG + TI} = \frac{TP + TN}{TP + FN + FP + TN} \quad 0 \leq A \leq 1$$

ERRO:

- Proporção de Classificações Incorretas para o total de elementos (Genuínos e Impostores).

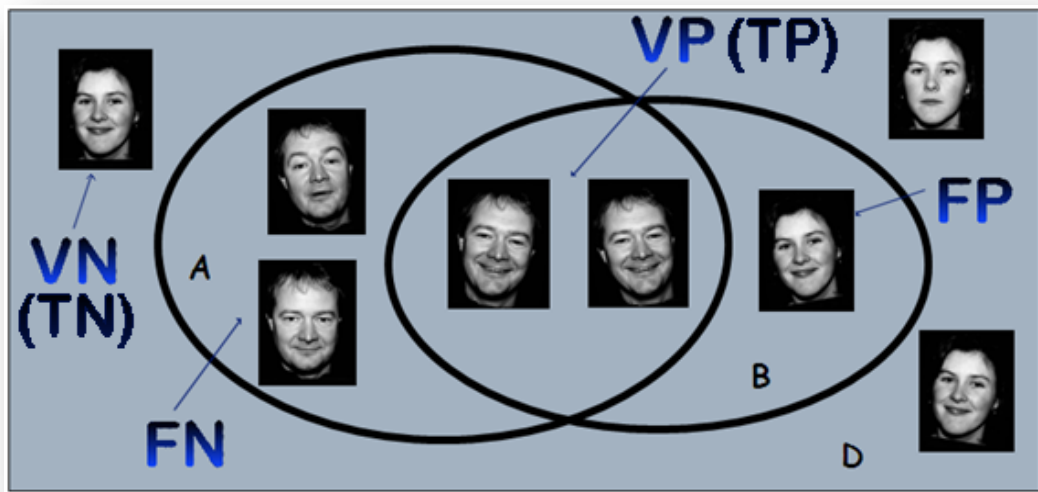
$$Err = \frac{TCI}{TS} = \frac{FP + FN}{TG + TI} = \frac{FP + FN}{TP + FN + FP + TN} = 1 - A$$

Saída\Entrada	Genuíno (Positivo)	Impostores (Negativos)	
Positivo	VP(TP)	FP	TSP(B)
Negativo	FN	VN(TN)	TSN
	TG(A)	TI	TS

Exemplo:

O conjunto A é uma classe, ou melhor, o
Total de Genuínos $\rightarrow TG = 4$

$$TG = TP + FN$$



O conjunto B é o
conjunto de elementos
mais semelhantes
retornados pelo
classificador, ou seja,
O Total de Saídas
Positivas $\rightarrow TSP=3$

$$TSP = TP + FP$$

O conjunto D é o conjunto dos
Impostores, ou seja, o Total de
Impostores $\rightarrow TI = 4$

$$TI = TN + FP$$

Saída\Entrada	Genuíno (Positivo)	Impostores (Negativos)	
Positivo	VP(TP)	FP	TSP(B)
Negativo	FN	VN(TN)	TSN
	TG(A)	TI	TS

Exemplo:

Sensibilidade(Recall): $TPR = R = \frac{TP}{TP+FN} = \frac{2}{2+2} = 0,5$ (True Positive Rate)

Razão de falso positivo: $FPR = \frac{FP}{TN+FP} = \frac{1}{3+1} = 0,25$ (False Positive Rate)

Razão de falso negativos: $FNR = \frac{FN}{TP+FN} = \frac{2}{2+2} = 0,5$ (False Negative Rate)

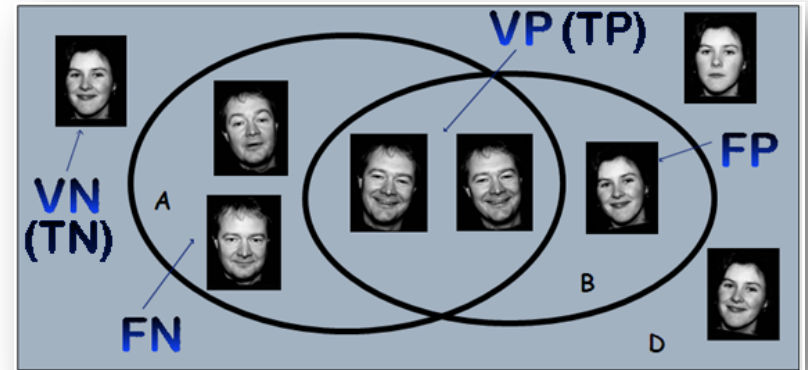
Especificidade: $TNR = E = \frac{TN}{TN+FP} = \frac{3}{3+1} = 0,75$

Valor previsto na saída positiva(Precisão): $P = \frac{TP}{TP+FP} = \frac{2}{2+1} = 0,67$

Valor previsto da saída negativa: $N = \frac{TN}{TSN} = \frac{TN}{TN+FN} = \frac{3}{3+2} = 0,6$

Acurácia: $A = \frac{TCC}{TS} = \frac{TP+TN}{TP+FN+FP+TN} = \frac{2+3}{2+2+1+3} = 0,625$

Erro: $Err = \frac{FP+FN}{TP+FN+FP+TN} = \frac{1+2}{2+2+1+3} = 0,375$



Considerações:

- Se o Classificador tiver um alto valor para Falsos Negativos (FN), ele não classifica corretamente quem deveria ser classificado, ou seja, sua SENSIBILIDADE é baixa.

$$TPR = R = \frac{TP}{TG} = \frac{TP}{TP + FN} \quad 0 \leq TPR \leq 1$$

- Se o Classificador tiver um alto valor de Falsos Positivos (FP), ele classifica elementos que não deveriam ser classificados, ou seja, sua ESPECIFICIDADE e sua PRECISÃO são baixas.

$$TNR = E = \frac{TN}{TI} = \frac{TN}{TN + FP} = 1 - FPR \quad 0 \leq E \leq 1$$

$$P = \frac{TP}{TSP} = \frac{TP}{TP + FP} \quad 0 \leq P \leq 1$$

Curvas ROC

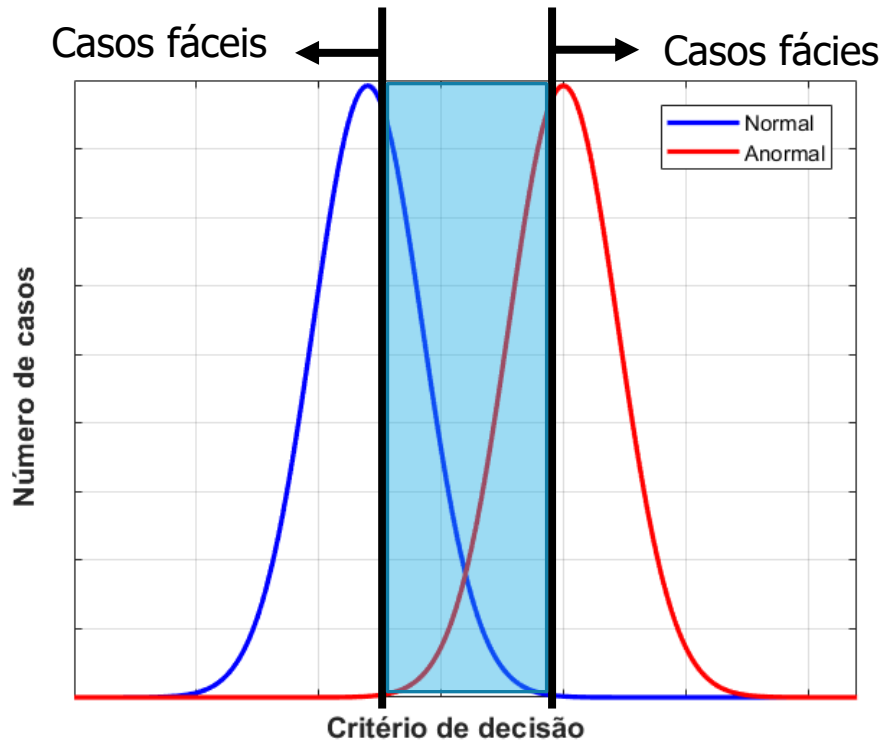
AVALIAÇÃO DE CLASSIFICADORES

Curva ROC (Característica de Operação do Receptor)

- **Exemplo:** Análise de exames pelo médico;
- Cada médico tem seu critério de decisão
- O paciente deve ser classificado em duas classes (diagnósticos):
- Diagnóstico:
 - Normal
 - Anormal
- Número de casos normais e anormais são representados por gaussianas dependendo do critério adotado pelo classificador.

Curva ROC (Característica de Operação do Receptor)

Análise do número de casos de pacientes normais e anormais

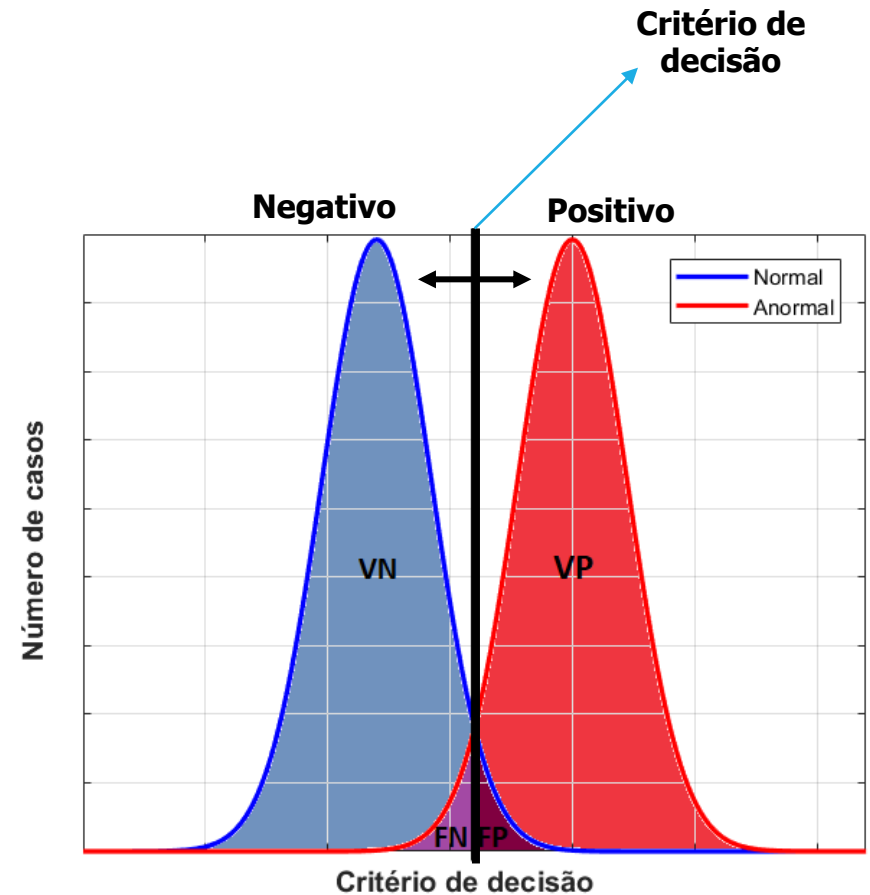
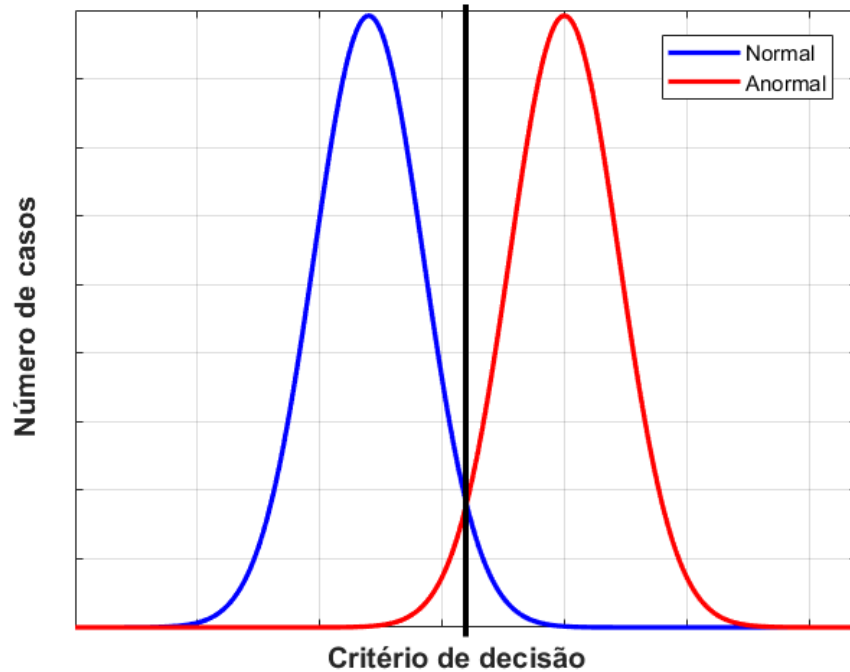


- Casos fáceis
- Casos difíceis
 - O que fazer com o meio?

Curva ROC (Característica de Operação do Receptor)

Diferentes análises de observadores:

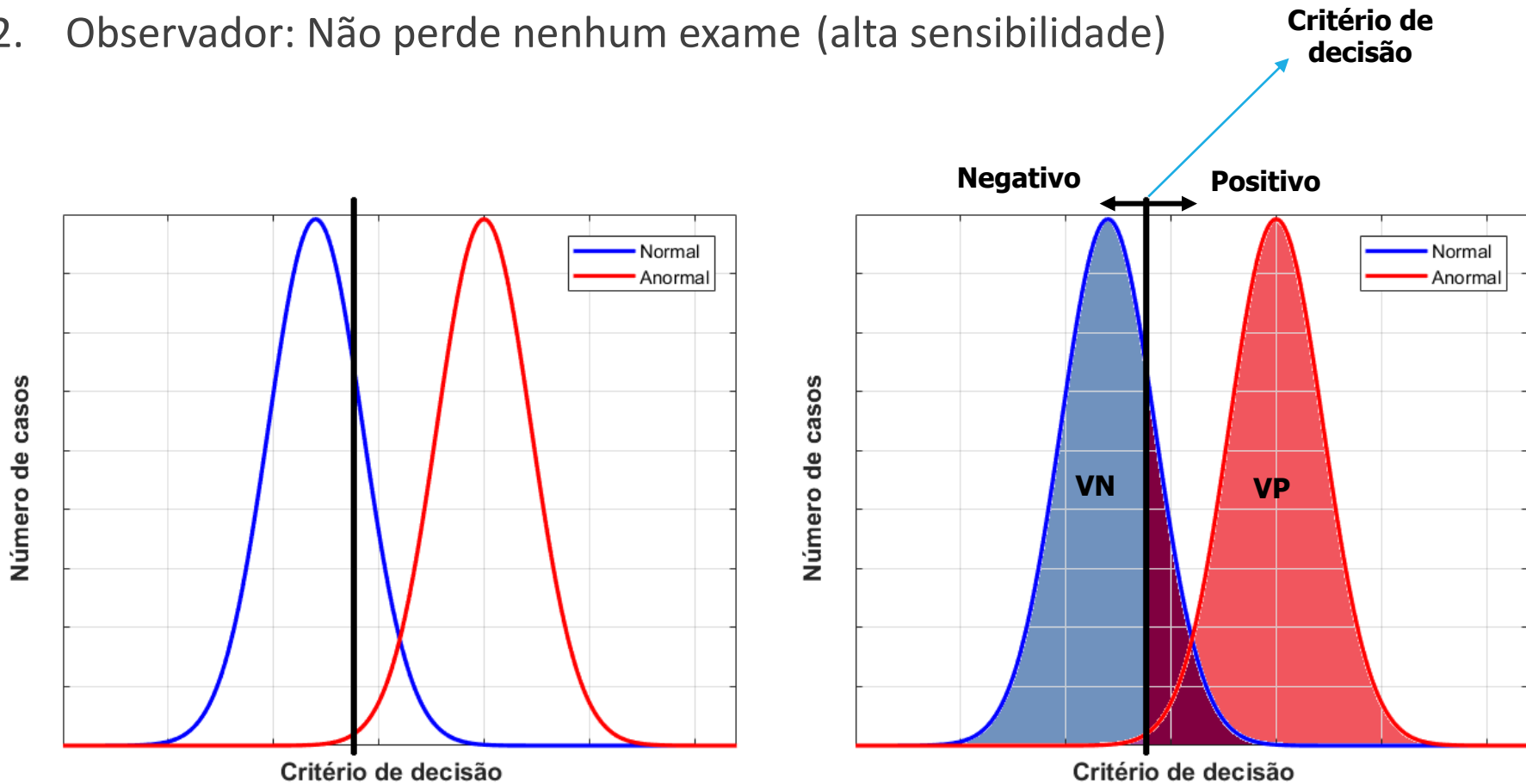
1. Observador usual:



Curva ROC (Característica de Operação do Receptor)

Diferentes análises de observadores:

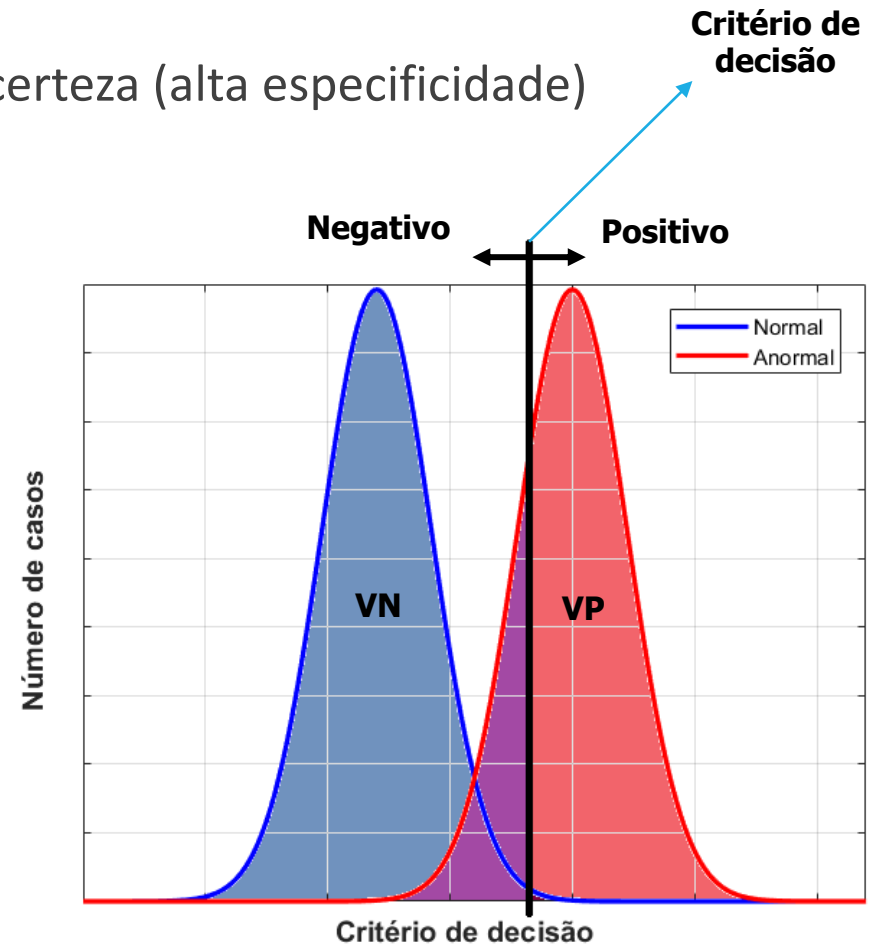
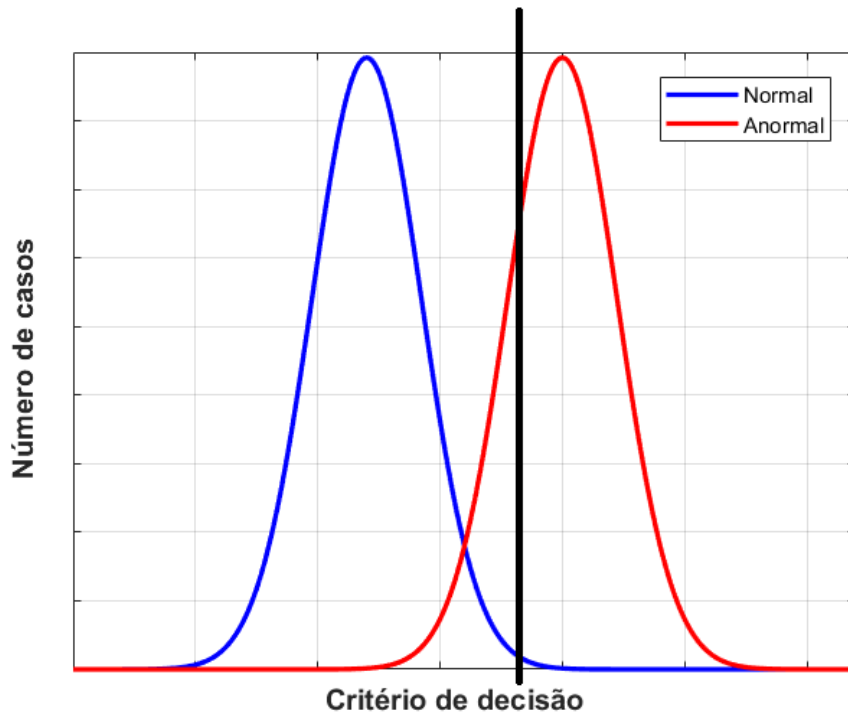
2. Observador: Não perde nenhum exame (alta sensibilidade)



Curva ROC (Característica de Operação do Receptor)

Diferentes análises de observadores:

3. Apenas confirma os exames que tem certeza (alta especificidade)

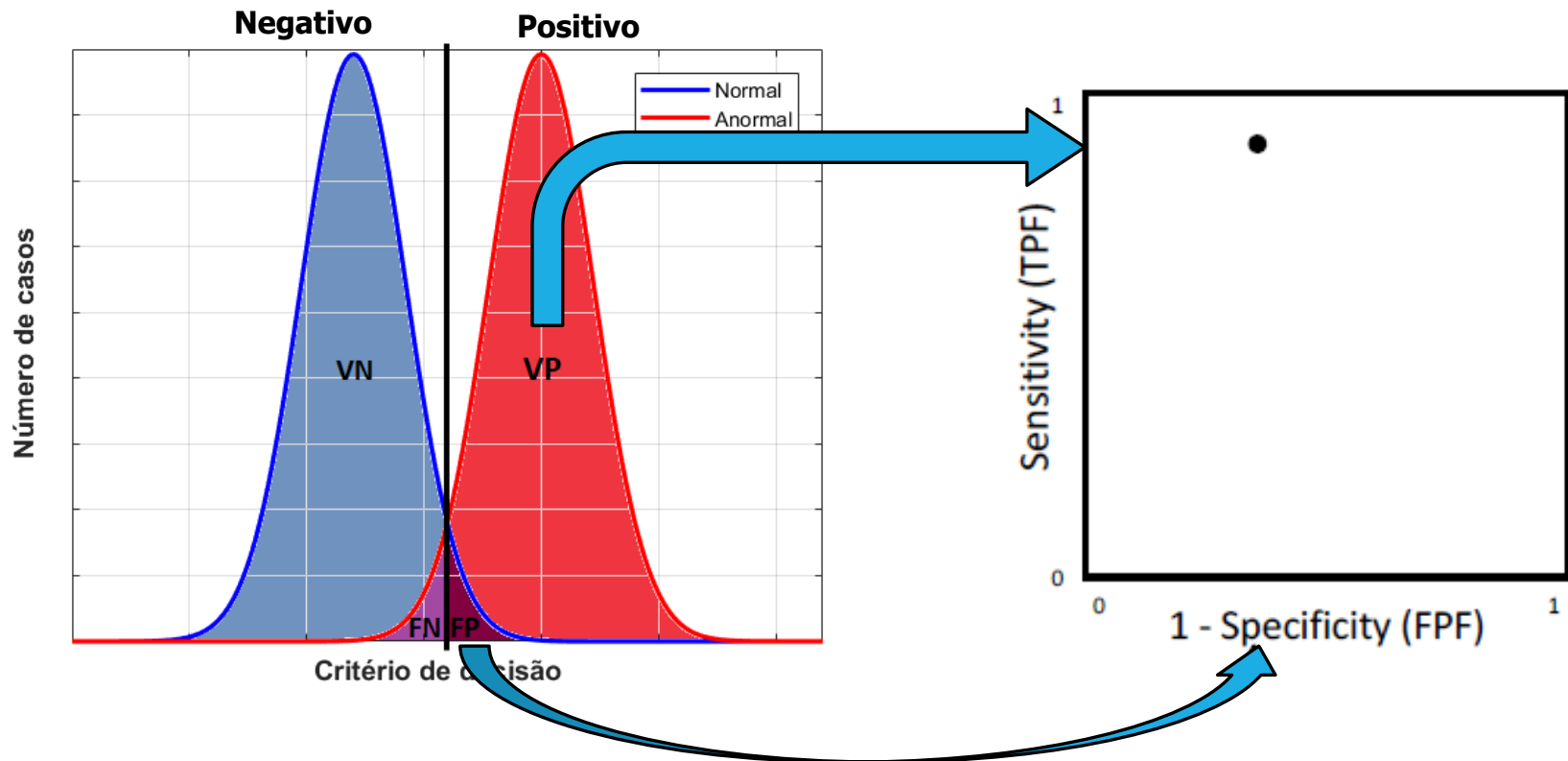


Curva ROC (Característica de Operação do Receptor)

A partir das diferentes respostas analisadas, é possível construir curva ROC do sistema

Gráfico:

- Sensibilidade vs. (1 – especificidade)

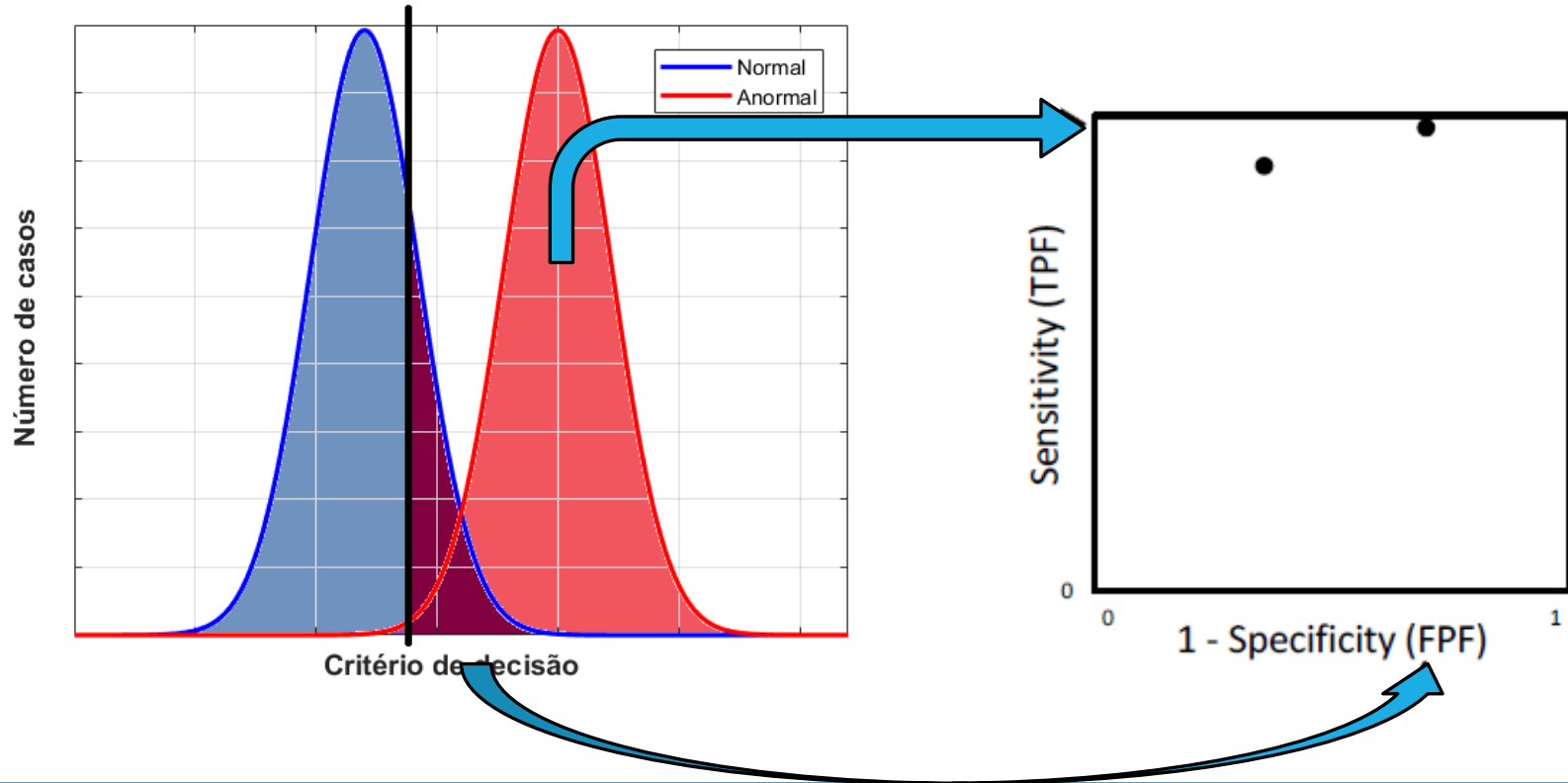


Curva ROC (Característica de Operação do Receptor)

A partir das diferentes respostas analisadas, é possível construir curva ROC do sistema

Gráfico:

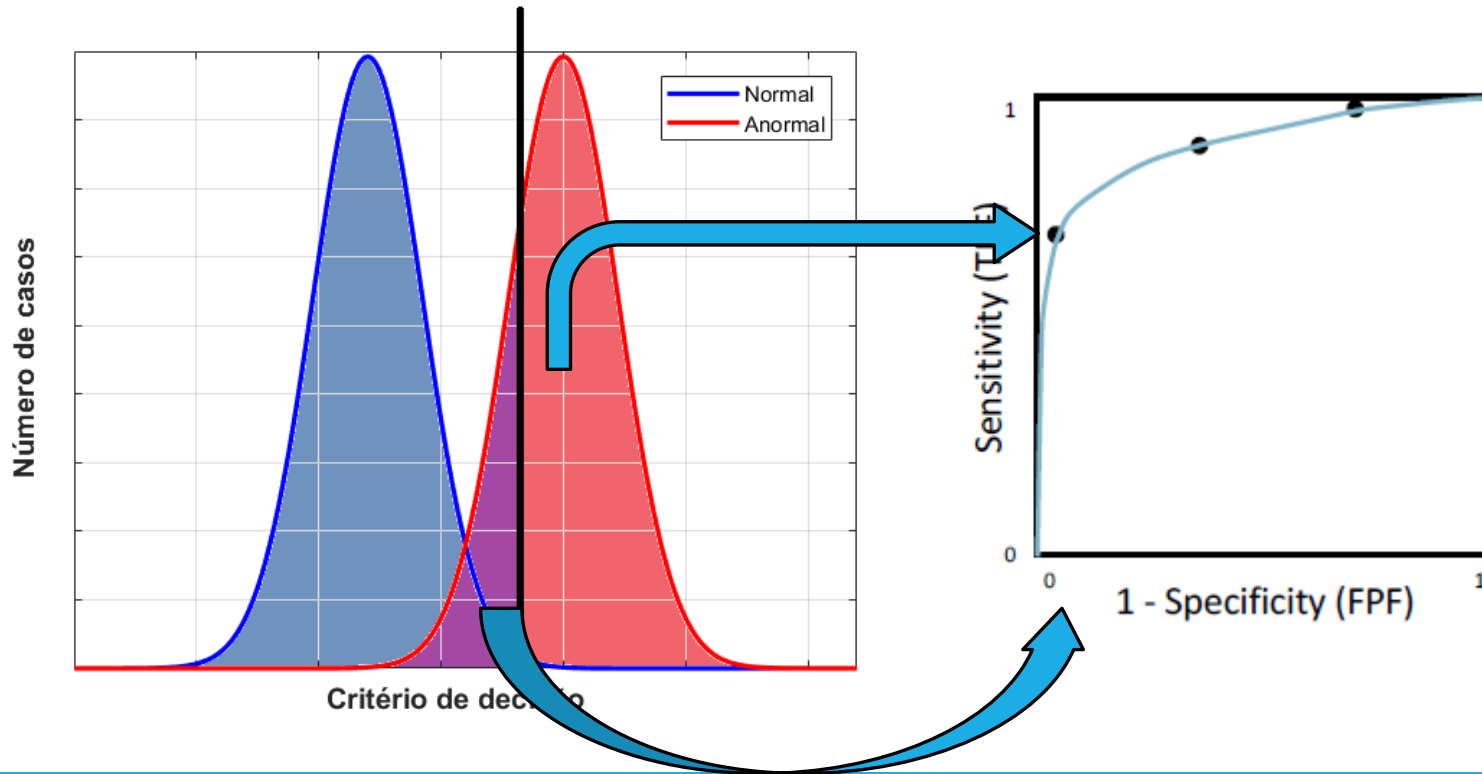
- Sensibilidade vs. (1 – especificidade)



Curva ROC (Característica de Operação do Receptor)

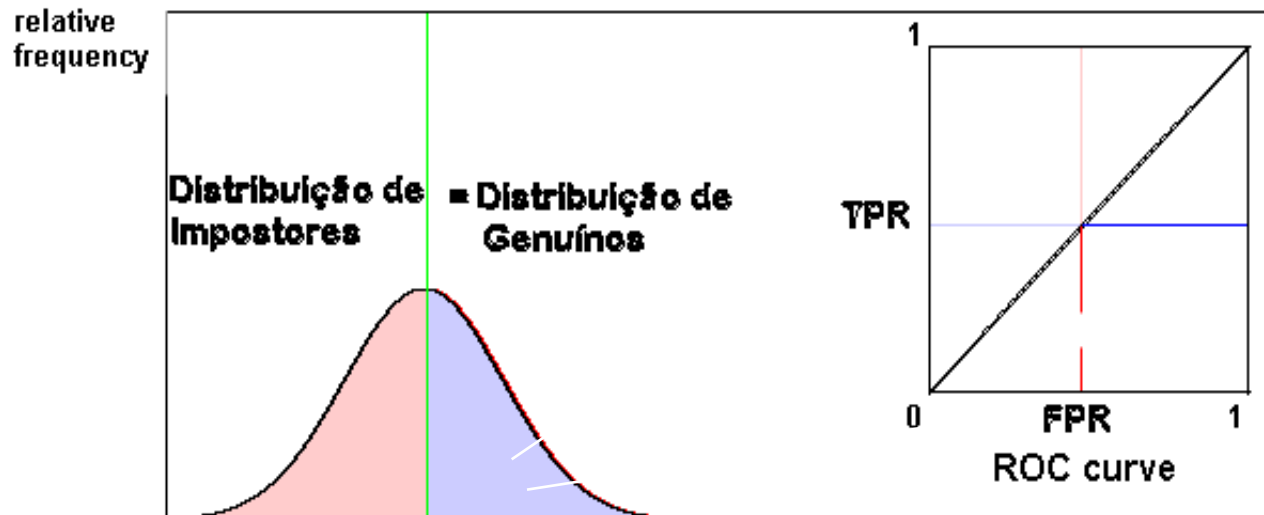
Gráfico:

- Sensibilidade vs. (1 – especificidade)
- Para cada threshold, calcula-se a quantidade de Verdadeiros Positivos (TP), Falsos Positivos (FP). Calcula-se a TPR (True Positive Rate) e a FPR (False Positive Rate) e constrói-se a Curva ROC.



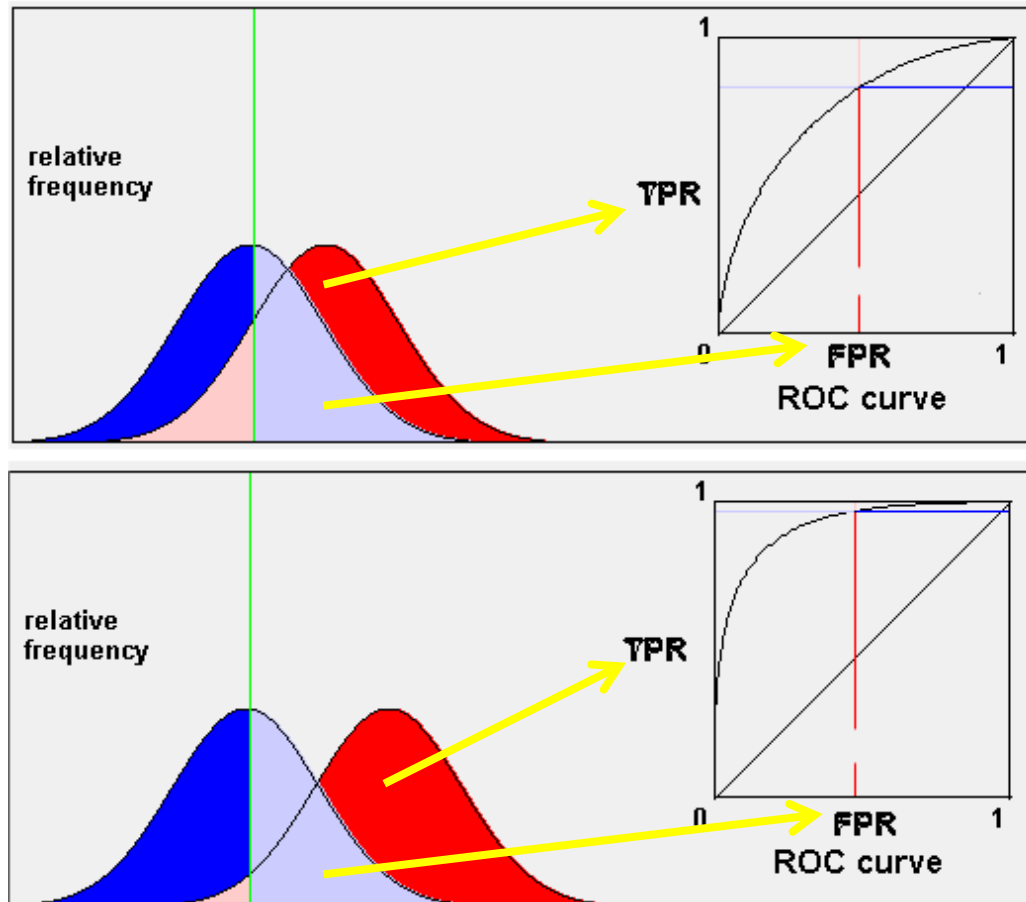
Discriminação das Classes:

- 1) Se o Classificador não consegue discriminar entre Impostores e Genuínos, qualquer ponto cai sobre a diagonal da Curva ROC (Random Guessing).



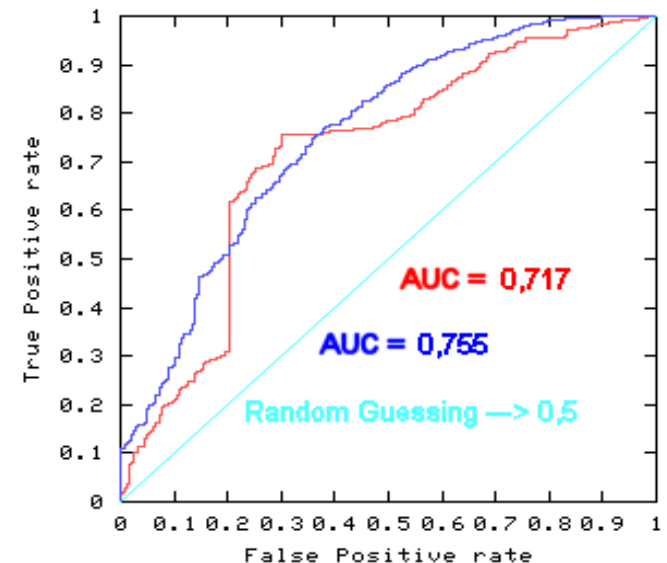
Discriminação das Classes:

- 2) Quanto mais discriminante é o Classificador, mais afastadas são as distribuições e a Curva ROC afasta-se da diagonal.



Área sob a curva ROC (AUC).

- A AUC informa o quanto o modelo é capaz de distinguir entre classes.
- O cálculo da área sob a curva ROC (AUC) permite comparar classificadores. O pior caso é $AUC=0,5$ para o Classificador Aleatório (Random Guessing).
- Quanto melhor o classificador $AUC \rightarrow 1$.
- $AUC < 0,5$ correspondem a sistemas com critério invertido



Matriz de Confusão (Confusion Matrix)

AVALIAÇÃO DAS CLASSIFICAÇÕES EM DIFERENTES
CLASSES

Matriz de Confusão

- Matriz de Confusão é uma matriz quadrada $n \times n$ onde as colunas representam as n classes reais de entrada e as n linhas as n classes de saída do classificador.
- Cada célula fornece o número de elementos classificados.
- Uma Matriz de Confusão mostra com mais detalhes a operação do classificador.
- É possível verificar quais as classes que o classificador confunde mais e tomar decisões que otimizem sua saída.
- Quando um conjunto de dados é não balanceado, ou seja, o número de elementos de cada classe é diferente, a acurácia não é uma boa métrica. O resultado do classificador é melhor analisado através de uma Matriz de Confusão.

Matriz de Confusão

		Entrada					FPR
		Classe 1	Classe 2	Classe 3	Classe 4	Classe n	
Saída	Classe 1	TP	FN2/FP1	FN3/FP1	FN4/FP1	FNn/FP1	FPR1
	Classe 2	FN1/FP2	TP	FN3/FP2	FN4/FP2	FNn/FP2	FPR2
	Classe 3	FN1/FP3	FN2/FP3	TP	FN4/FP3	FNn/FP3	FPR3
	Classe 4	FN1/FP4	FN2/FP4	FN3/FP4	TP	FNn/FP4	FPR4
	Classe n	FN1/FPn	FN2/FPn	FN3/FPn	FN4/FPn	TP	FPRn
FNR		FNR1	FNR2	FNR3	FNR4	FNRn	TPR

$$P = \frac{TP}{TSP} = \frac{TP}{TP+FP} \text{ (Precisão)}$$

$$FNR = \frac{FN}{TG} = \frac{FN}{TP+FN} \text{ (Falsa Rejeição)}$$

$$TPR = R = \frac{TP}{TG} = \frac{TP}{TP+FN} \text{ (Sensibilidade)}$$

$$FPR = \frac{FP}{TI} = \frac{FP}{TN+FP} \text{ (Falsa Aceitação)}$$

Exemplo:

4 Classes:

- Classe 1 = 100 cachorros
- Classe 2 = 50 coelhos
- Classe 3 = 80 gatos
- Classe 4 = 120 lobos



		Entrada				FPR
		Cachorros	Coelhos	Gatos	Lobos	
Saída	Cachorros	80	0	2	10	0,046
	Coelhos	10	50	8	5	0,071
	Gatos	0	0	70	5	0,018
	Lobos	10	0	0	100	0,042
FNR		0,2	0	0,125	0,167	0,857

		Entrada				FPR
		Cachorros	Coelhos	Gatos	Lobos	
Saída	Cachorros	80	0	2	10	0,046
	Coelhos	10	50	8	5	0,071
	Gatos	0	0	70	5	0,018
	Lobos	10	0	0	100	0,042
FNR		0,2	0	0,125	0,167	0,857

Total de Genuínos (Total de Elementos) = $100 + 50 + 80 + 120 = 350$

Verdadeiros Negativos:

- TN_1 = Todos os que NÃO são cachorros = 250 (350-100)
- TN_2 = Todos os que NÃO são coelhos = 300 (350-50)
- TN_3 = Todos os que NÃO são gatos = 270 (350-80)
- TN_4 = Todos os que NÃO são lobos = 230 (350-120)

Falsa Aceitação → Falsos Positivos em uma Classe

$$FPR_n = \frac{FP}{TI} = \frac{FP}{TN + FP} = \frac{\sum_1^4 FP_I}{TN_n + \sum_1^4 FP_I}$$

		Entrada				FPR
		Cachorros	Coelhos	Gatos	Lobos	
Saída	Cachorros	80	0	2	10	0,046
	Coelhos	10	50	8	5	0,071
	Gatos	0	0	70	5	0,018
	Lobos	10	0	0	100	0,042
FNR		0,2	0	0,125	0,167	0,857

Para a Classe 1 → Razão de elementos que foram classificados como cachorro sem ser.

$$FPR_1 = \frac{2 + 10}{250 + 2 + 10} = \frac{12}{262} = 0,046$$

Para a Classe 2 → Razão de elementos que foram classificados como coelho sem ser.

$$FPR_2 = \frac{10 + 8 + 5}{300 + 10 + 8 + 5} = \frac{23}{323} = 0,071$$

Falsa Aceitação → Falsos Positivos em uma Classe

$$FPR_n = \frac{FP}{TI} = \frac{FP}{TN + FP} = \frac{\sum_1^4 FP_I}{TN_n + \sum_1^4 FP_I}$$

		Entrada				FPR
		Cachorros	Coelhos	Gatos	Lobos	
Saída	Cachorros	80	0	2	10	0,046
	Coelhos	10	50	8	5	0,071
	Gatos	0	0	70	5	0,018
	Lobos	10	0	0	100	0,042
FNR		0,2	0	0,125	0,167	0,857

Para a Classe 3 → Razão de elementos que foram classificados como gato sem ser.

$$FPR_3 = \frac{5}{270 + 5} = \frac{5}{275} = 0,018$$

Para a Classe 4 → Razão de elementos que foram classificados como lobo sem ser.

$$FPR_4 = \frac{10}{230 + 10} = \frac{10}{240} = 0,042$$

Falsa Rejeição → Falsos Negativos em uma Classe

$$FNR_n = \frac{FN}{TG} = \frac{FN}{TP + FN} = \frac{\sum_1^4 FN_I}{TP_n + \sum_1^4 FN_I}$$

		Entrada				FPR
		Cachorros	Coelhos	Gatos	Lobos	
Saída	Cachorros	80	0	2	10	0,046
	Coelhos	10	50	8	5	0,071
	Gatos	0	0	70	5	0,018
	Lobos	10	0	0	100	0,042
FNR		0,2	0	0,125	0,167	0,857

Para a Classe 1 → Razão de cachorros não classificados como cachorros.

$$FNR_1 = \frac{10 + 10}{80 + 10 + 10} = \frac{20}{100} = 0,2$$

Para a Classe 2 → Razão de coelhos não classificados como coelhos.

$$FNR_2 = \frac{0}{50 + 0} = 0$$

Falsa Rejeição → Falsos Negativos em uma Classe

$$FNR_n = \frac{FN}{TG} = \frac{FN}{TP + FN} = \frac{\sum_1^4 FN_I}{TP_n + \sum_1^4 FN_I}$$

		Entrada				FPR
		Cachorros	Coelhos	Gatos	Lobos	
Saída	Cachorros	80	0	2	10	0,046
	Coelhos	10	50	8	5	0,071
	Gatos	0	0	70	5	0,018
	Lobos	10	0	0	100	0,042
FNR		0,2	0	0,125	0,167	0,857

Para a Classe 3 → Razão de gatos não classificados como gatos.

$$FNR_3 = \frac{8 + 2}{70 + 8 + 2} = \frac{10}{80} = 0,125$$

Para a Classe 4 → Razão de lobos não classificados como lobos.

$$FNR_4 = \frac{10 + 5 + 5}{100 + 10 + 5 + 5} = \frac{20}{120} = 0,167$$

SENSIBILIDADE → Recall → Probabilidade de uma Amostra relevante ser classificada corretamente.

$$TPR = R = \frac{TP}{TG} = \frac{TP}{TP + FN} = \frac{80 + 50 + 70 + 100}{350} = \frac{300}{350} = 0,857$$

PRECISÃO → Probabilidade de uma Amostra classificada ser relevante.

$$P = \frac{TP}{TSP} = \frac{TP}{TP + FP} = \frac{300}{350} = 0,857$$

		Entrada				FPR
		Cachorros	Coelhos	Gatos	Lobos	
Saída	Cachorros	80	0	2	10	0,046
	Coelhos	10	50	8	5	0,071
	Gatos	0	0	70	5	0,018
	Lobos	10	0	0	100	0,042
FNR		0,2	0	0,125	0,167	0,857

Diz-se, então que este classificador tem Precisão, ou Índice de Classificações Corretas de 85,7 %.

$$P = \frac{TP}{TSP} = \frac{TP}{TP + FP} = \frac{300}{350} = 0,857$$

Erro: Taxa de elementos classificados erroneamente.

$$Err = \frac{TCI}{TS} = \frac{FP + FN}{TG + TI} = \frac{FP + FN}{TP + FN + FP + TN} = 1 - A = 0,143$$

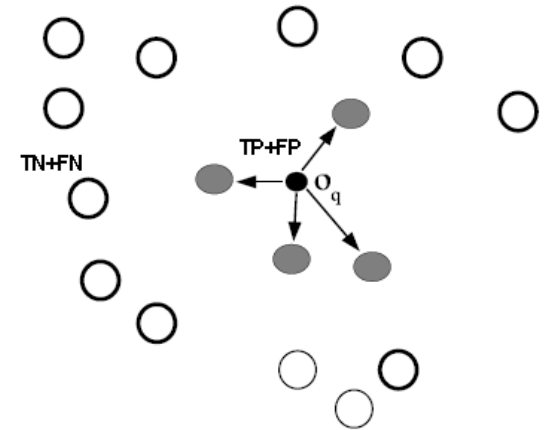
		Entrada				FPR
		Cachorros	Coelhos	Gatos	Lobos	
Saída	Cachorros	80	0	2	10	0,046
	Coelhos	10	50	8	5	0,071
	Gatos	0	0	70	5	0,018
	Lobos	10	0	0	100	0,042
FNR		0,2	0	0,125	0,167	0,857

Curvas Recall x Precision

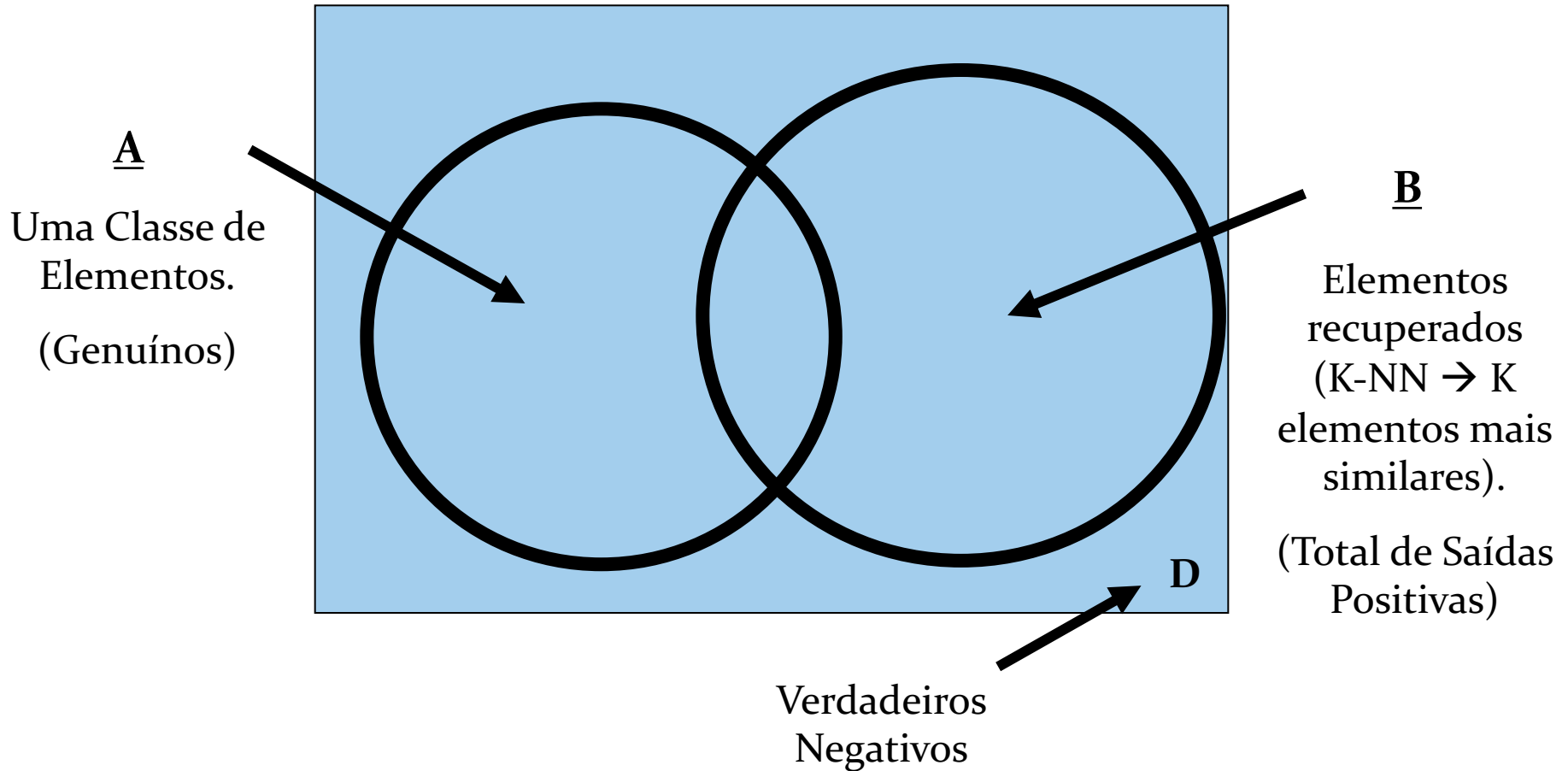
Avaliação de classificadores baseado na
Relevância.

Exemplo: Classificação pelos K Vizinhos Mais Próximos (K-NN).

- Os K elementos mais semelhantes ao elemento O_q , são classificados como Genuínos (TP + FP).
- Variando-se o valor de K (threshold), controla-se o número de Genuínos classificados.
- Quanto menor o valor de K (maior threshold) mais “rígido”, ou Conservador, é o classificador e quanto maior é o valor de K, mais elementos são classificados como Genuínos (menor é o threshold) e o classificador é dito mais “tolerante” ou Liberal.



A informação relevante.

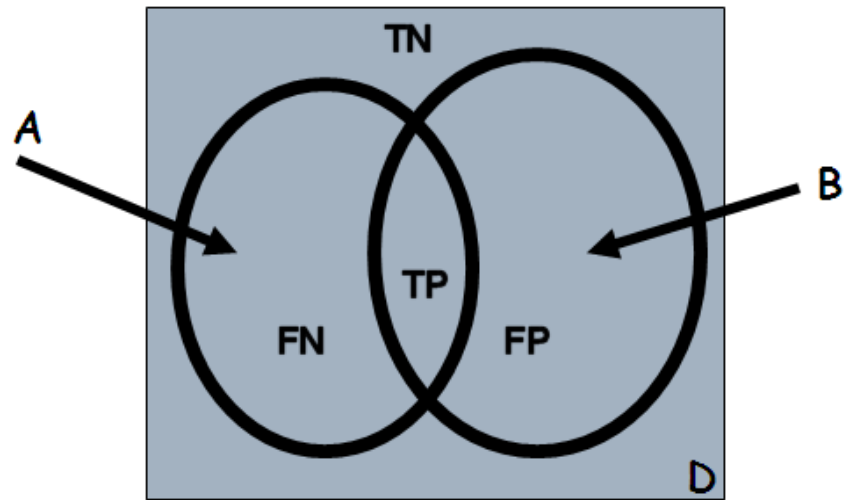


Medidas baseadas na Relevância: Recall x Precision (Relembrando)

Recall: Probabilidade de uma Amostra relevante ser recuperada.

$$recall = \frac{|A \cap B|}{|A|} = \frac{TP}{TP + FN}$$

$$precision = \frac{|A \cap B|}{|B|} = \frac{TP}{TP + FP}$$



Precision: Probabilidade de uma Amostra recuperada ser relevante

Exemplo:

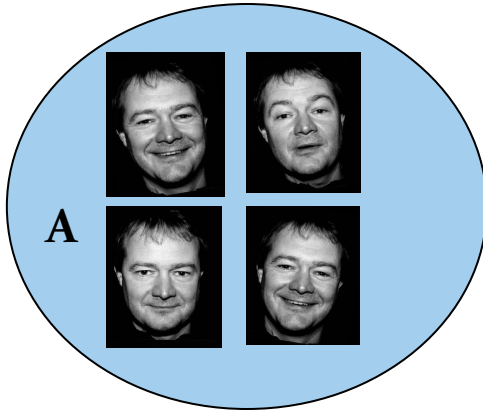
- Seja A = a classe de Elementos representada pelas imagens:

Figura 1

Figura 2

Figura 3

Figura 4



E o conjunto de Imagens da
Base = 8 Elementos

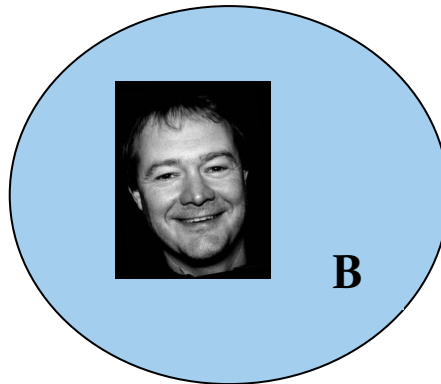


Considerando-se a Figura 1.

- Seja a imagem a ser classificada (Imagem de busca) a Figura 1:

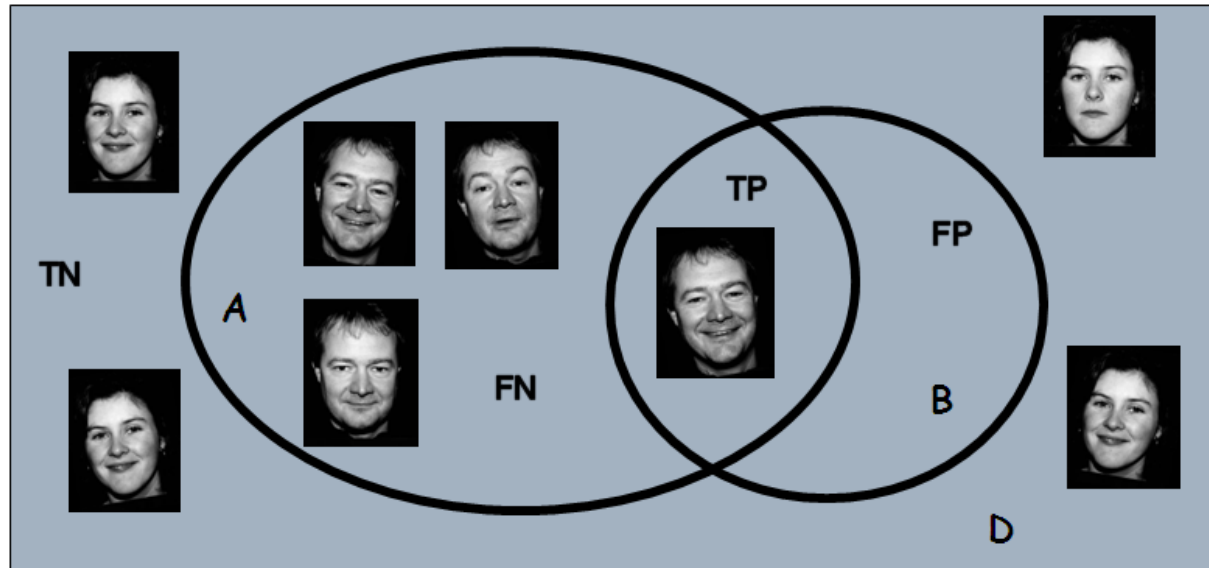


- Para $K=1$, ou seja, a 1ª imagem mais similar à Figura 1, retornada é:





□ Para $K=1$, o classificador retornou uma imagem mais similar (conjunto B):

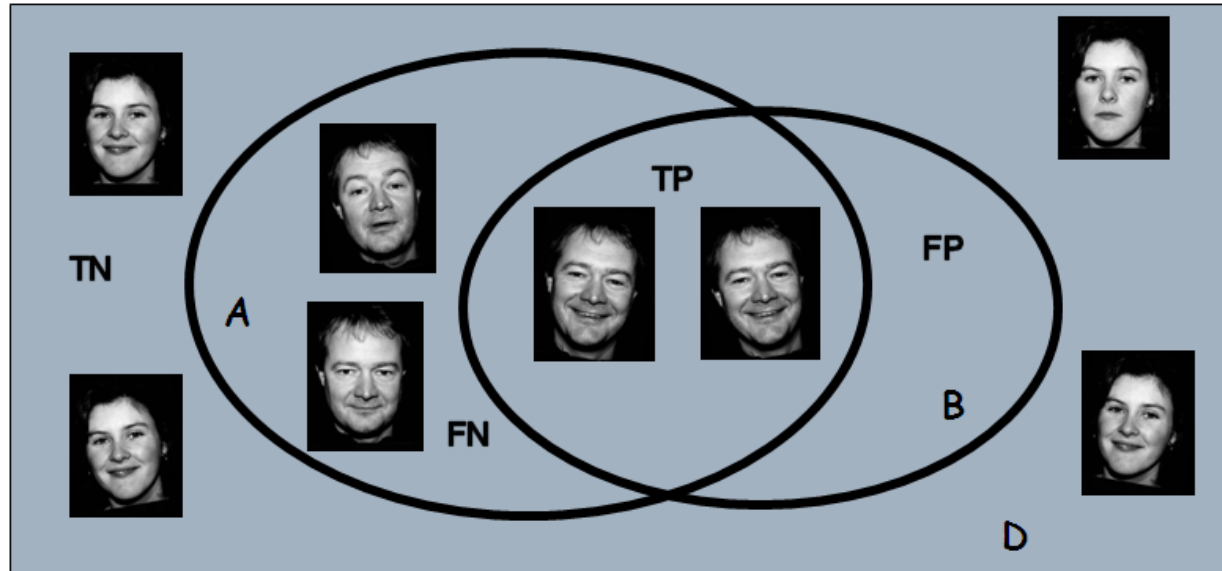


$$recall(1) = \frac{|A \cap B|}{|A|} = \frac{TP}{TP + FN} = \frac{1}{4} = 0,25$$

$$precision(1) = \frac{|A \cap B|}{|B|} = \frac{TP}{TP + FP} = \frac{1}{1} = 1$$



- Para $K=2$, o classificador retornou duas imagens mais similares (conjunto B):

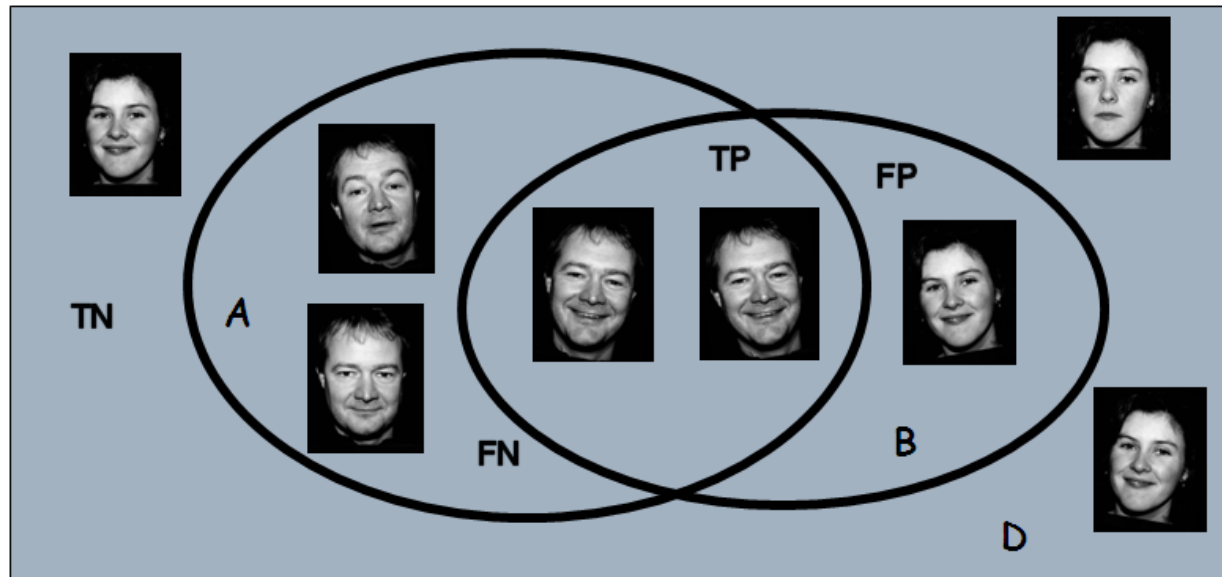


$$recall(2) = \frac{|A \cap B|}{|A|} = \frac{TP}{TP + FN} = \frac{2}{4} = 0,5$$

$$precision(2) = \frac{|A \cap B|}{|B|} = \frac{TP}{TP + FP} = \frac{2}{2} = 1$$



- Para $K=3$, o classificador retornou três imagens mais similares (conjunto B):

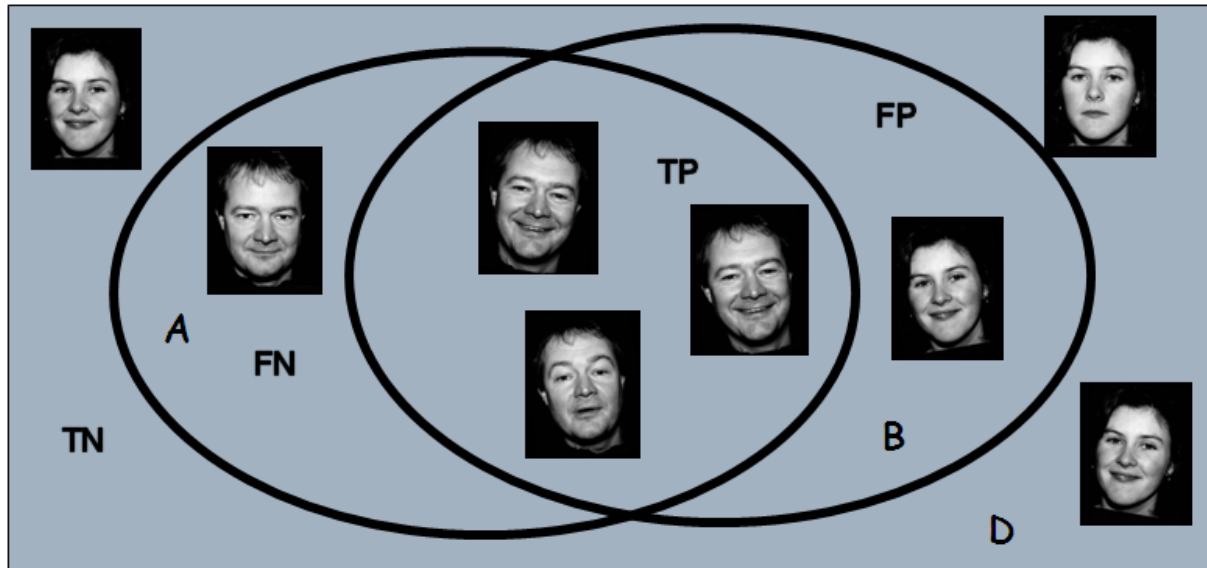


$$recall(3) = \frac{|A \cap B|}{|A|} = \frac{TP}{TP + FN} = \frac{2}{4} = 0,50$$

$$precision(3) = \frac{|A \cap B|}{|B|} = \frac{TP}{TP + FP} = \frac{2}{3} = 0,67$$



- Para $K=4$, o classificador retornou quatro imagens mais similares (conjunto B):



$$recall(4) = \frac{|A \cap B|}{|A|} = \frac{TP}{TP + FN} = \frac{3}{4} = 0,75$$

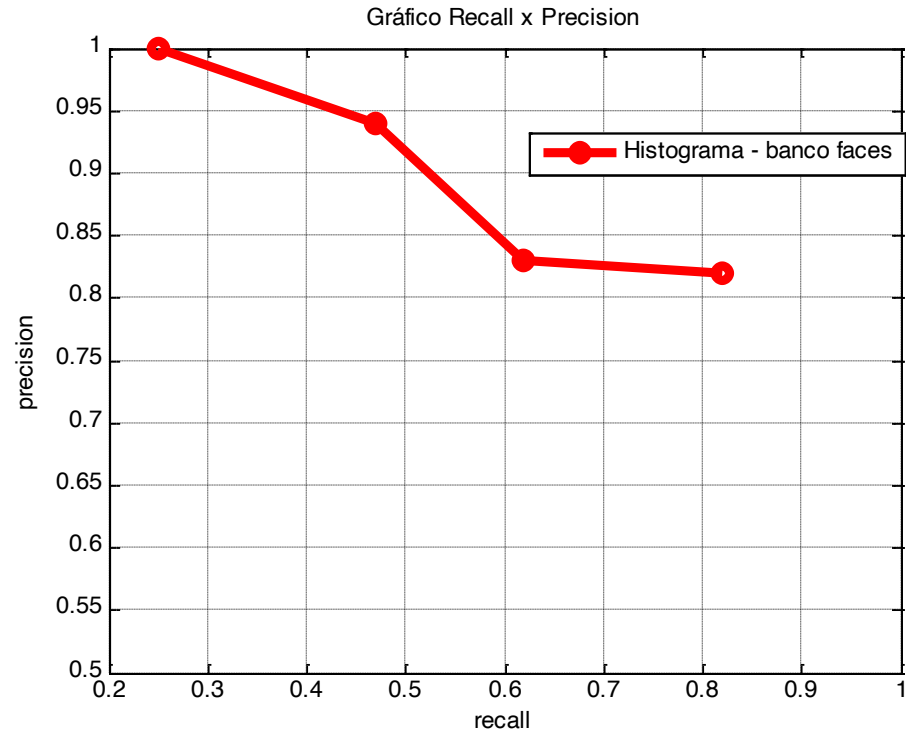
$$precision(4) = \frac{|A \cap B|}{|B|} = \frac{TP}{TP + FP} = \frac{3}{4} = 0,75$$

Calculando-se o par RP para cada Figura da base, tem-se:

Busca	K=1		K=2		K=3		K=4	
	Recall	Precision	Recall	Precision	Recall	Precision	Recall	Precision
Fig. 1	0,25	1,00	0,50	1,00	0,50	0,67	0,75	0,75
Fig. 2	0,25	1,00	0,50	1,00	0,75	1,00	1,00	1,00
Fig. 3	0,25	1,00	0,50	1,00	0,75	1,00	0,75	0,75
Fig. 4	0,25	1,00	0,50	1,00	0,50	0,67	0,75	0,75
Fig. 5	0,25	1,00	0,50	1,00	0,75	1,00	1,00	1,00
Fig. 6	0,25	1,00	0,50	1,00	0,75	1,00	0,75	0,75
Fig. 7	0,25	1,00	0,25	0,50	0,50	0,67	0,75	0,75
Fig. 8	0,25	1,00	0,50	1,00	0,50	0,67	0,75	0,75
Média	0,25	1,00	0,47	0,94	0,62	0,83	0,82	0,82

Curva Recall x Precision

❑ Quanto mais próximo do topo do gráfico, melhor é o desempenho do classificador.



❑ Para se comparar dois gráficos de classificadores (ou algoritmos) diferentes, alinham-se os pontos recuperados (mesmo valor de K) verticalmente, adotando-se o mesmo Recall para os dois e calculando-se as diferentes Precisões.

FIM