



Inatel

Modelos com imagens

Prof. Eng. Ranyeri do Lago Rocha
e-mail ranyeri.rocha@inatel.br

Inatel

Agenda

- Modelos com imagens

Todo o conteúdo deste documento está relacionado a direito autoral e é de circulação restrita, porquanto de propriedade exclusiva da Fundação Instituto Nacional de Telecomunicações (CNPJ 24.492.886/0001-04), protegido por força das disposições da Lei n.º 9.610/1998. A utilização deste material sem prévia e expressa autorização da proprietária constituirá infração à lei, com repercussões tanto na esfera civil quanto criminal.

Modelos com imagens

Modelos com imagens

- Tarefas
 - Detecção
 - Segmentação
 - Classificação
 - *Pose estimation*
 - OBB
- Tracking
- Ajuste fino de um modelo
- Análise de desempenho: Matriz de confusão, erro e curva ROC

Detecção

Detecção é a tarefa que envolve detectar objetos em uma imagem ou *frame* de vídeo e desenhar *bouding boxes* ao seu entorno.

Os objetos detectados são classificados em diferentes categorias

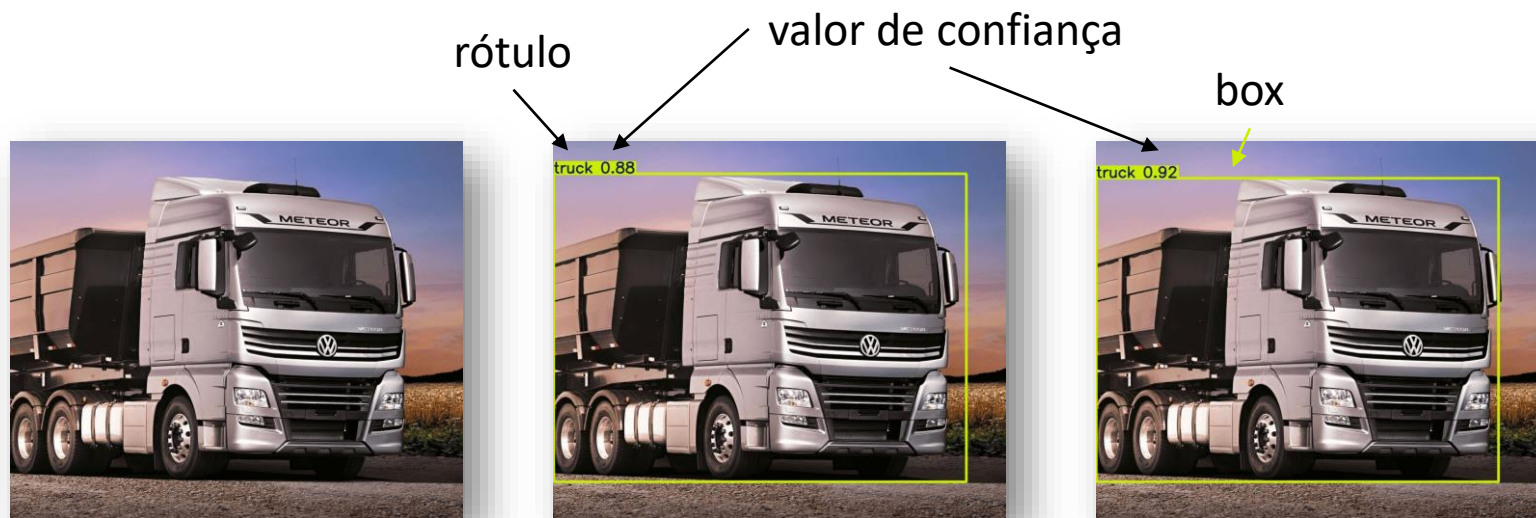
Os modelos podem detectar objetos únicos ou múltiplos em uma única image ou *frame*

Entrada: imagem

Saída: *bouding boxes* para cada objeto, o rótulo associado e o valor de confiança para cada *box*

Detecção

Detecção é a tarefa que envolve detectar objetos em uma imagem ou *frame* de vídeo e desenhar *bouding boxes* ao seu entorno.



Detecção

Detecção é a tarefa que envolve detectar objetos em uma imagem ou *frame* de vídeo e desenhar *bouding boxes* ao seu entorno.

Quando usar?

Identificação de objetos de interesse sem a necessidade de conhecer onde o objeto está exatamente ou o seu formato.



Segmentação

Segmentação é a tarefa que envolve segmentar uma imagem em diferentes regiões baseado no seu conteúdo. A cada região é associado um rótulo baseado em seu conteúdo.

Vai além da detecção. Identifica os objetos e segmenta-os do resto da imagem

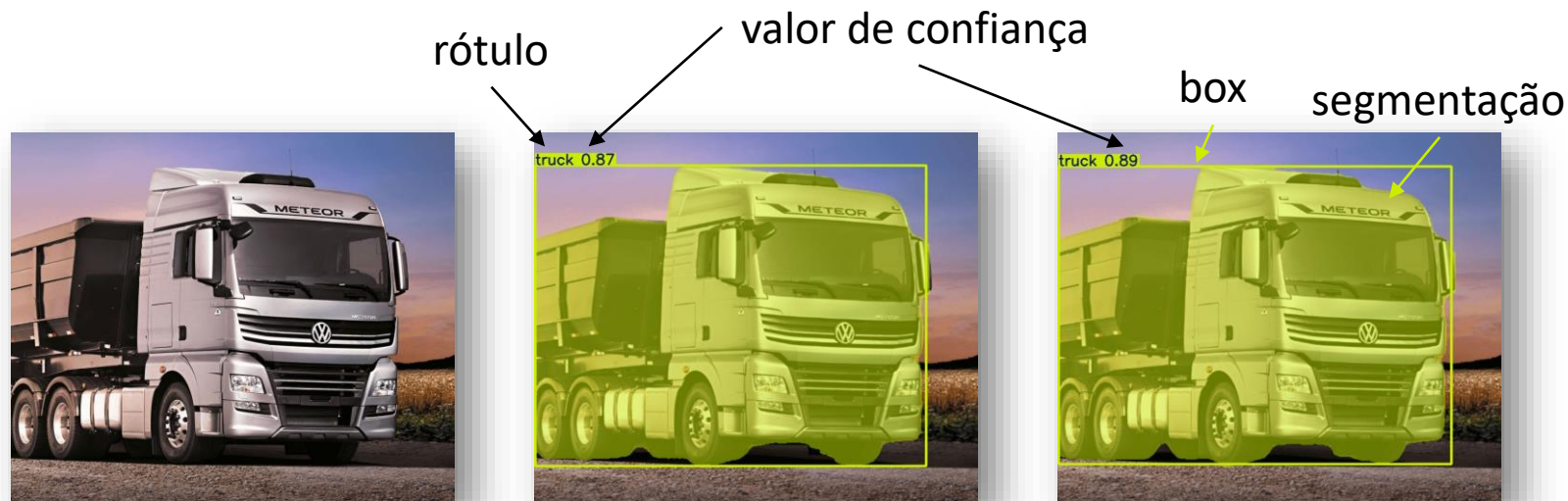
Os modelos podem segmentar objetos únicos ou múltiplos em uma única image ou *frame*

Entrada: imagem

Saída: conjunto de máscaras ou contornos que delimita cada objeto, um rótulo associado a uma classe e os valores de confiança

Segmentação

Segmentação é a tarefa que envolve segmentar uma imagem em diferentes regiões baseado no seu conteúdo. A cada região é associado um rótulo baseado em seu conteúdo.

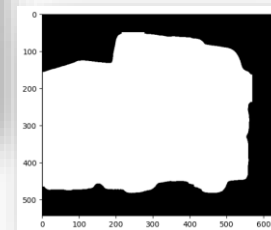


Segmentação

Segmentação é a tarefa que envolve segmentar uma imagem em diferentes regiões baseado no seu conteúdo. A cada região é associado um rótulo baseado em seu conteúdo.

Quando usar?

Segmentação é útil quando é necessário saber não apenas onde o objeto está em uma imagem, mas também, seu formato exato.



Classificação

Classificação é uma tarefa que envolve classificar uma imagem em diferentes categorias. É a tarefa mais simples entre detecção, segmentação e classificação.

A classificação é feita na imagem inteira, não apenas em partes dela

Entrada: imagem

Saída: rótulo único e o valor de confiança

Classificação

Classificação é uma tarefa que envolve classificar uma imagem em diferentes categorias. É a tarefa mais simples entre detecção, segmentação e classificação.



Classificação

Classificação é uma tarefa que envolve classificar uma imagem em diferentes categorias. É a tarefa mais simples entre detecção, segmentação e classificação.

Quando usar?

Útil quando precisa apenas da classe de pertencimento da imagem.



Pose

Pose Estimation ou detecção de pontos chave/pose envolve detectar pontos específicos em uma imagem ou *frame* de vídeo.

Os pontos detectados (keypoints) são usados para rastreamento de movimento ou estimação de pose.
Localização dos pontos em uma imagem, como junções, *landmarks* ou característica importante.

Entrada: imagem

Saída: conjunto de pontos que representa os *keypoints* de um objeto na imagem, rótulo e confiança de cada ponto

Pose

Pose Estimation ou detecção de pontos chave/pose envolve detectar pontos específicos em uma imagem ou *frame* de vídeo.



OBB – *Oriented Bouding Box*

Oriented Bounding Box está associado ao processo de detecção de objetos em imagens, porém com informação extra de angulação para localizar objetos de forma mais acurada.

O box é ajustado exatamente aos objetos na imagem.

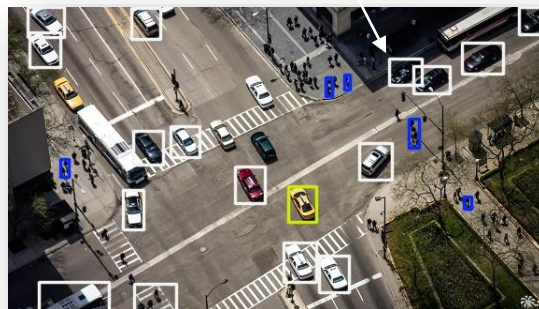
Entrada: imagem

Saída: conjunto de *bounding box* rotacionados, rótulo de classe e valor de confiança

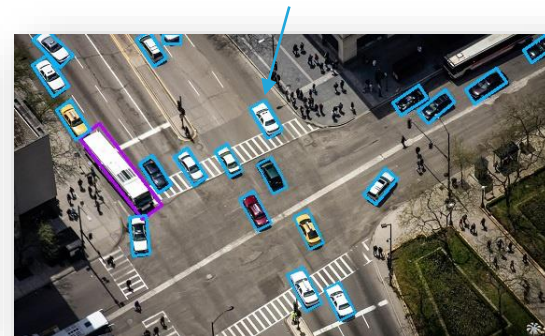
OBB – *O*riented *B*ouding *B*ox

Oriented Bounding Box está associado ao processo de detecção de objetos em imagens, porém com informação extra de angulação para localizar objetos de forma mais acurada.

box



oriented box



Tracking

Tracking ou rastreamento de objetos é uma tarefa crítica para manutenção da identificação do objeto detectado, enquanto ele estiver na cena.

O principal desafio é manter o ID único em casos drásticos de oclusão dos objetos.
Os modelos estão aptos a *tracking* de objetos únicos ou múltiplos.

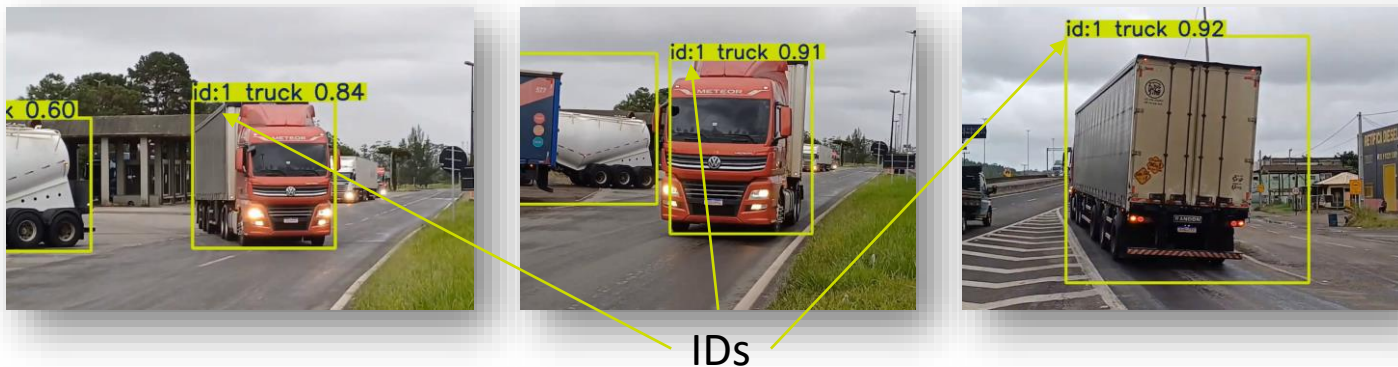
Entrada: sequência de imagens/frames

Saída: saídas consistentes com as de detecção de objetos, com a informação extra do ID do objeto

Tracking

Tracking ou rastreamento de objetos é uma tarefa crítica para manutenção da identificação do objeto detectado, enquanto ele estiver na cena.

Sequência de frames



Tracking

Tracking ou rastreamento de objetos é uma tarefa crítica para manutenção da identificação do objeto detectado, enquanto ele estiver na cena.

Quando usar?

Não há limitações para *tracking*. Aplicações como monitoramento de segurança, análise em tempo real, são comuns em diferentes contextos.



Ajuste fino de um modelo

Ajuste fino de um modelo, ou transferência de aprendizado (*Transfer Learning*) permite usar o conhecimento adquirido de um modelo para resolver um problema diferente.

O modelo não é treinado *from scratch*

Otimização de tempo, recurso e, frequentemente, melhorando o desempenho

Por que utilizar?

Transferir o aprendizado, utilizar as características aprendidas em um conjunto de dados grande, em um conjunto menor. “Conjunto de dados grande” é geralmente um conjunto massivo, como ImageNet.

ImageNet: 1000 classes
1.281.167 imagens para treinamento
50.000 para validação
100.000 para teste.

Ajuste fino de um modelo

Ajuste fino de um modelo, ou transferência de aprendizado (*Transfer Learning*) permite usar o conhecimento adquirido de um modelo para resolver um problema diferente.

O modelo não é treinado *from scratch*

Otimização de tempo, recurso e, frequentemente, melhorando o desempenho

Onde utilizar?

O conceito pode ser aplicado em várias áreas. Entretanto é particularmente usado para visão computacional e processamento de linguagem natural.

Ajuste fino de um modelo

Ajuste fino de um modelo, ou transferência de aprendizado (*Transfer Learning*) permite usar o conhecimento adquirido de um modelo para resolver um problema diferente.

Etapa 1 - Pré-treinamento

Treinamento envolve todas as camadas

Etapa 2 – Fine Tuning

Treinamento envolve, geralmente, somente as camadas finais.

O que é aprendido nas camadas iniciais de uma rede convolucional?

Ajuste fino de um modelo

Ajuste fino de um modelo, ou transferência de aprendizado (*Transfer Learning*) permite usar o conhecimento adquirido de um modelo para resolver um problema diferente.

O processo de transferência de aprendizado:

1. Reduz o tempo de treinamento
2. Necessita de menos dados para treinamento
3. Possui uma generalização melhorada
4. Otimiza o processo de implementação e desenvolvimento de modelos em produção

Análise de desempenho: Matriz de confusão, erro e curva ROC

Matriz de Confusão

A matriz de confusão é geralmente utilizada para dar maiores informações sobre o desempenho do classificador.

- A matriz de confusão contabiliza o número de classificações corretas e incorretas para cada uma das K classes existentes.

Parâmetros:

$$C = \begin{bmatrix} C_{ii} & C_{ij} \\ C_{ji} & C_{jj} \end{bmatrix}$$

Sendo:

C_{ij} indica quantos padrões da classe i foram designados à classe j

C_{ji} indica quantos padrões da classe j foram designados à classe i

Na diagonal principal, temos o número de classificações corretas

Análise de desempenho: Matriz de confusão, erro e curva ROC

Matriz de Confusão

TP – True Positive: Número de exemplos da classe positiva classificados corretamente

TN – True Negative: Número de exemplos da classe negativa classificados corretamente

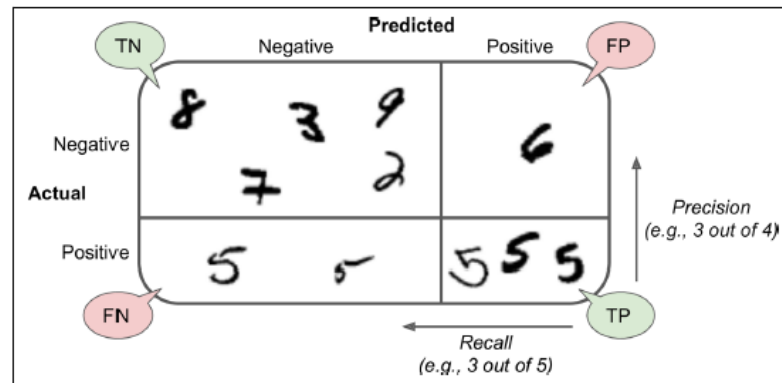
FP – False Positive: Número de exemplos da classe negativa classificados como positivos

FN – False Negative: Número de exemplo da classe positiva classificados como negativos

		Predicted		
		Negative	Positive	
Actual	Negative	8 3 9	6	FP
	Positive	5 5	5 5 5	TP

Precision (e.g., 3 out of 4)

Recall (e.g., 3 out of 5)

A confusion matrix diagram with handwritten numbers. The matrix is a 2x2 grid. The top row is labeled 'Negative' and the bottom row is labeled 'Positive' under the 'Actual' column. The left column is labeled 'Negative' and the right column is labeled 'Positive' under the 'Predicted' column. The top-left cell contains the numbers '8 3 9'. The top-right cell contains the number '6'. The bottom-left cell contains the numbers '5 5'. The bottom-right cell contains the numbers '5 5 5'. There are four callout boxes: 'TN' (green) pointing to the top-left cell, 'FP' (red) pointing to the top-right cell, 'FN' (red) pointing to the bottom-left cell, and 'TP' (green) pointing to the bottom-right cell. To the right of the matrix, an upward arrow is labeled 'Precision (e.g., 3 out of 4)'. Below the matrix, a leftward arrow is labeled 'Recall (e.g., 3 out of 5)'.

N_+ é o número de padrões pertencentes à classe positiva = $TP + FN$

N_- é o número de padrões pertencentes à classe negativa = $FP + TN$

N é o número total de padrões = $TP + FN + FP + TN$

Análise de desempenho: Matriz de confusão, erro e curva ROC

Matriz de Confusão

Algumas métricas mais concisas podem ser mais interessantes:

Acurácia:

$$acurácia = \frac{TP + TN}{N}$$

Precisão (acurácia das predições positivas):

$$precision = \frac{TP}{TP + FP}$$

Recall (taxa das instâncias positivas que foram corretamente detectadas):

$$recall, TPR, sensitivity = \frac{TP}{TP + FN}$$

A confusion matrix diagram illustrating classification performance using handwritten digits. The matrix is divided into four quadrants: True Negative (TN), False Positive (FP), False Negative (FN), and True Positive (TP). The top row represents the 'Predicted' class (Negative/Positive), and the bottom row represents the 'Actual' class (Negative/Positive). The left column represents the 'Actual' class (Negative/Positive), and the right column represents the 'Predicted' class (Negative/Positive). The matrix contains handwritten digits: 8, 3, 9, 6, 7, 2, 5, 5, 5. The matrix is annotated with labels: TN (True Negative) in the top-left corner, FP (False Positive) in the top-right corner, FN (False Negative) in the bottom-left corner, and TP (True Positive) in the bottom-right corner. The matrix is also annotated with 'Precision (e.g., 3 out of 4)' and 'Recall (e.g., 3 out of 5)'.

	Negative	Positive
Negative	8 3 9	6
Positive	5 5 5	5 5 5

Labels: TN (True Negative), FP (False Positive), FN (False Negative), TP (True Positive).
Precision (e.g., 3 out of 4)
Recall (e.g., 3 out of 5)

Análise de desempenho: Matriz de confusão, erro e curva ROC

Matriz de Confusão

Algumas métricas mais concisas podem ser mais interessantes:

Taxa de Falso Negativo (exemplo positivos classificados incorretamente):

$$FNR = \frac{FN}{TP + FN}$$

Taxa de Falso Positivo (exemplo negativos classificados incorretamente):

$$FPR = \frac{FP}{TN + FP}$$

Taxa de erro:

$$erro = \frac{FP + FN}{N}$$

	Negative	Positive
Negative	8 3 9	6
Positive	5 5	5 5 5

Labels around the matrix: TN (top-left), FP (top-right), FN (bottom-left), TP (bottom-right).
Annotations: Precision (e.g., 3 out of 4) pointing to TP; Recall (e.g., 3 out of 5) pointing to TP.

Análise de desempenho: Matriz de confusão, erro e curva ROC

Matriz de Confusão

Algumas métricas mais concisas podem ser mais interessantes:

F1-Score (média harmônica de precisão e *recall*):

Especialmente útil para classes com distribuição não balanceada.

	Predicted	
	Negative	Positive
Negative Actual	8 7	9 2
Positive Actual	5	5 5 5

Labels: TN (True Negative), FP (False Positive), FN (False Negative), TP (True Positive).
Precision (e.g., 3 out of 4) is indicated for the Positive Predicted group.
Recall (e.g., 3 out of 5) is indicated for the Positive Actual group.

$$F1\ score = \frac{2}{\frac{1}{precision} + \frac{1}{recall}} = \frac{TP}{TP + \frac{FN + FP}{2}}$$

Análise de desempenho: Matriz de confusão, erro e curva ROC

Matriz de Confusão

Exemplo:

$$\begin{aligned}
 \text{acurácia} &= \frac{TP + TN}{N} = \frac{3 + 5}{11} = 0,72 \\
 \text{erro} &= \frac{FP + FN}{N} = \frac{1 + 2}{11} = 0,27 \\
 \text{precision} &= \frac{TP}{TP + FP} = \frac{3}{3 + 1} = 0,75 \\
 \text{recall, TPR, sensitivity} &= \frac{TP}{TP + FN} = \frac{3}{3 + 2} = 0,6 \\
 \text{FNR} &= \frac{FN}{TP + FN} = \frac{2}{3 + 2} = 0,4 \\
 \text{FPR} &= \frac{FP}{TN + FP} = \frac{1}{5 + 1} = 0,16 \\
 \text{F1 score} &= \frac{TP}{TP + \frac{FN + FP}{2}} = \frac{3}{3 + \frac{2 + 1}{2}} = 0,66
 \end{aligned}$$

		Predicted		
		Negative	Positive	
Actual	Negative	<div>8</div> <div>3</div> <div>9</div> <div>7</div> <div>2</div>	<div>6</div>	<div>Precision (e.g., 3 out of 4)</div>
	Positive	<div>5</div> <div>5</div> <div>5</div> <div>5</div> <div>5</div>	<div>5</div> <div>5</div> <div>5</div>	
		<div>TN</div>	<div>FP</div>	
		<div>FN</div>	<div>TP</div>	<div>Recall (e.g., 3 out of 5)</div>

Análise de desempenho: Matriz de confusão, erro e curva ROC

Matriz de Confusão

Condiserações:

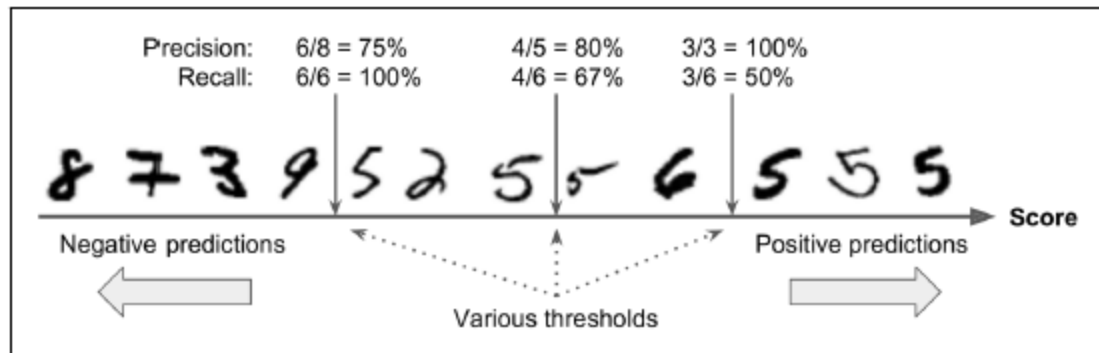
- Se um classificador tiver um alto valor para FN, ele não classifica corretamente o que deveria ser classificado, ou seja, sua **sensibilidade** é BAIXA.
- Se um classificador teve um alto valor de FP, ele classifica elementos que não deveriam ser classificados, ou seja, sua especificidade e a sua **precisão** são BAIXAS.
- **Precisão = 1** significa que, para uma determinada classe, cada exemplo classificado realmente pertence a ela. *Entretanto, isso não dá informações a respeito de quantas amostras desta classe foram classificadas de forma incorreta.*
- **Recall = 1** indica que todos os exemplos de uma classe foram classificados corretamente. *Porém, isso não traz informações a respeito de quantos exemplos associados a outras classes foram classificados como esta mesma classe.*

Análise de desempenho: Matriz de confusão, erro e curva ROC

Matriz de Confusão

Condiserações:

- O tradeoff Recall x Precision



- Recall e Precision costumam ser analisadas em conjunto. O *F-score*, como já apresentado, é uma média harmônica ponderada entre estas duas métricas. Quando o fator de ponderação é 1 (um), tem-se o *F1-score*. Assim, valores de *F1* próximos de 1 indicam que o classificar tem uma boa performance, tanto na **Precision** quanto no **Recall**

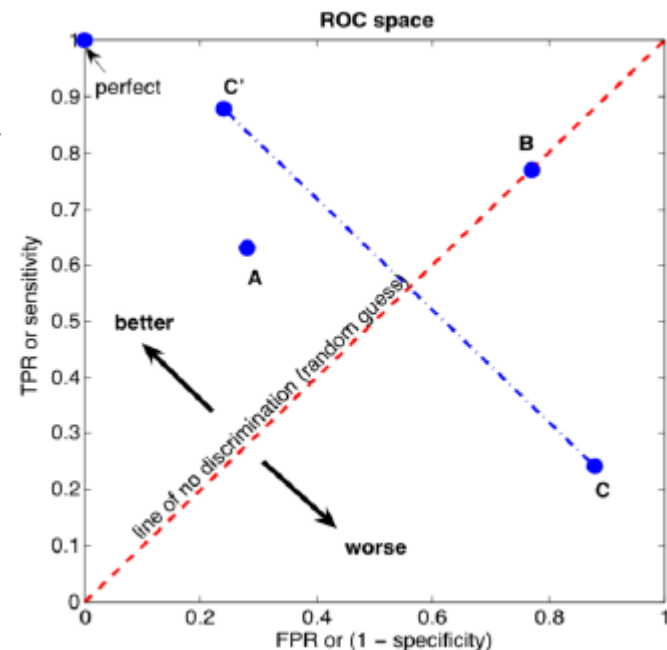
Análise de desempenho: Matriz de confusão, erro e curva ROC

Curva ROC

A curva ROC (Receiver Operating Characteristic) é outra ferramenta muito comum para classificadores binários (ou as versões OvO ou OvA).

Basicamente, a curva ROC é um gráfico 2D com o eixo vertical definido como TPR (True Positive Rate) e o eixo horizontal definido como FPR (False Positive Rate).

- Os pontos apresentados no gráficos, como, A, B, C, cada um apresenta uma análise diferente do modelo.
- O ponto marcado como “perfect” é o modelo perfeito, ideal

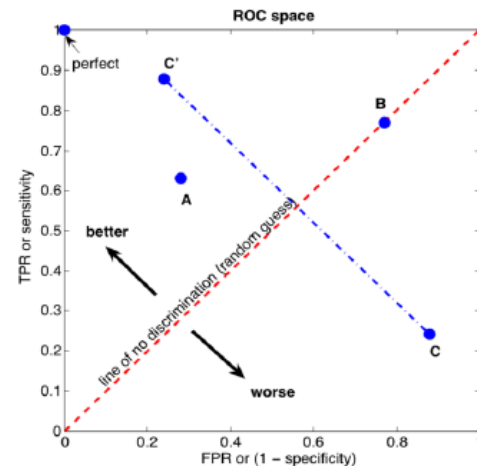


Análise de desempenho: Matriz de confusão, erro e curva ROC

Curva ROC

Algumas análises importantes:

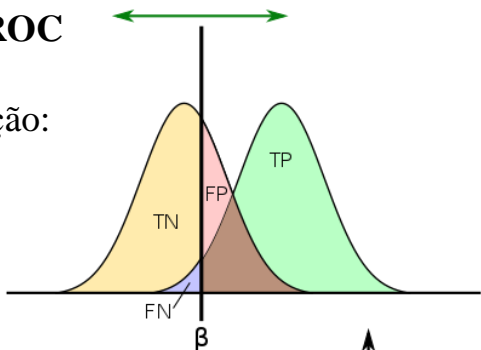
- Quanto maior o recall (TPR), mais falsos positivos (FPR) o classificador produz
- Um bom classificador está tão longe da **linha aleatória possível**
- **A linha diagonal está associada a um classificador aleatório.**
- Quanto mais a direita na curva, menos seletivo é o modelo.
- Quanto mais a esquerda na curva, mais seletivo é o modelo.
- O objetivo é encontrar o valor ideal para o projeto, assim como na análise Recall x Precision



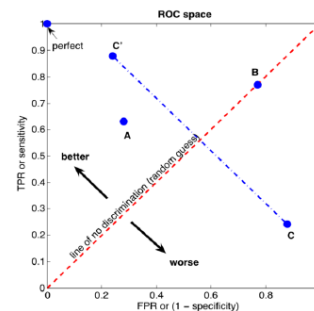
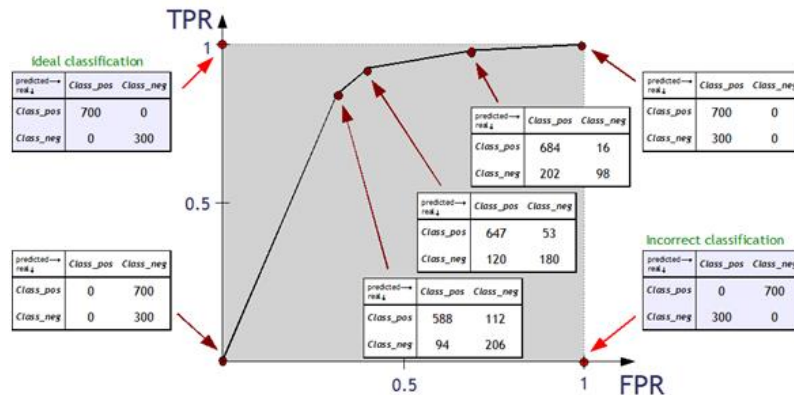
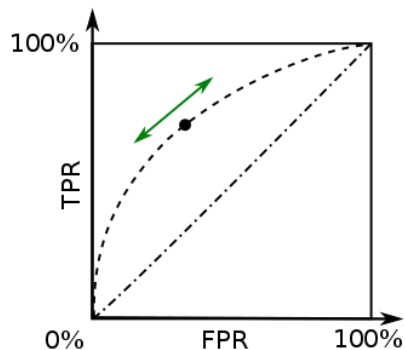
Análise de desempenho: Matriz de confusão, erro e curva ROC

Curva ROC

Construção:



TP	FP
FN	TN



Análise de desempenho: Matriz de confusão, erro e curva ROC

Curva ROC

Área sobre a curva (AUC)

Analisar a curva dá vários insights importantes, mas, para efeito de comparação entre modelos é importante uma métrica, um valor único que traduza o comportamento da curva.

O ROC AUC (ROC Area Under the Curve) é este valor:

- Se igual a 1, temos um classificador perfeito
- Se igual a 0,5, temos um classificador aleatório

A análise é simples: o modelo com maior valor de ROC AUC é comparativamente melhor.

Análise de desempenho: Matriz de confusão, erro e curva ROC

Curva ROC

Vantagens:

- Análise de diferentes métricas de desempenho independente do limiar
- Auxílio ao estudo de diferentes limiares para lidar com problemas de classes com tamanhos discrepantes

Desvantagens:

- Adequada para problemas de classificação binária
- Para multi classes, deve-se utilizar a estratégia OvA e plotar várias curvas ROC

Análise de desempenho: Mean Average Precision

mAP – mean Average Precision

Mede de forma balanceada, *precision* e *recall*, fornecendo um bom entendimento de como o modelo identifica e localiza objetos em imagens e vídeos.

Relembre os conceitos:

Precision: se um modelo tem uma precisão alta, indica que os objetos encontrados tem grande chance de estarem corretos.

Recall: se uma modelo tem um recall alto, indica que o modelo é bom em encontrar a maioria dos objetos presentes na imagem.

Análise de desempenho: Mean Average Precision

mAP – mean Average Precision

Desta forma, mAP sintetiza as duas métricas calculando a precisão média entre diferentes níveis de recall.

1. Para cada categoria de objetos, faz-se um ranking dos valores de confiança
2. *Precision* e *Recall* são computados para vários valores de limiares de confiança (*Precision-Recall Curve*)
3. *Average Precision* é computada como área sobre a curva
4. mAP é obtido pela média dos valores de AP entre todas as classes

Inatel



inatel



inateloficial



ascominatel



inatel.tecnologias



company/inatel

Inatel

Inatel - Instituto Nacional de Telecomunicações
Campus em Santa Rita do Sapucaí - MG - Brasil
Av. João de Camargo, 510 - Centro - 37540-000
+55 (35) 3471 9200

Escritório em São Paulo - SP - Brasil
WTC Tower, 18º andar - Conjunto 1811/1812
Av. das Nações Unidas, 12.551 - Brooklin Novo - 04578-903
+55 (11) 3043 6015