

FAST – Bootcamp Machine Learning

Experimentos e
Avaliação de Classificadores

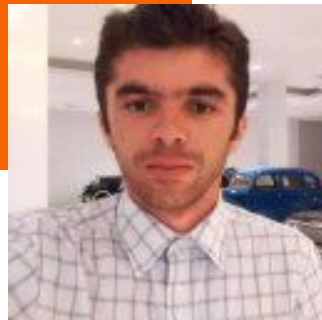
Pedro D. Marrero Fernandez
pdmf@cesar.school





Pedro D. Marrero

Pedro é formado em Ciência da Computação e com Mestrado pela Universidade de Oriente, Cuba, e Doutorado em Ciência da Computação pela Universidade Federal de Pernambuco (UFPE), Brasil. Colabora com diversos laboratórios de pesquisa internacional como: CAMBIA Lab, CALTECH, USA; Unitat de Gràfics i Visió per Ordinador i Intel·ligència Artificial, UIB, Espanha e VIISAR Lab, UFPE, Brasil. Suas principais contribuições e experiências são nas áreas de Computer Vision, Machine Learning, Deep Learning and Computational Photography.



Antônio J. Pinheiro

Antônio Janael Pinheiro é engenheiro de software do Centro de Estudos e Sistemas Avançados do Recife (CESAR). Ele recebeu graduação em Redes de Computadores pela Universidade Federal do Ceará em 2013, mestrado e doutorado em Ciência da Computação pela Universidade Federal de Pernambuco, em 2016 e 2020, respectivamente, e especialização em inteligência computacional aplicada pela Universidade Federal Rural de Pernambuco em 2021.



Blenda Guedes

Blenda é Cientista de Dados no C.E.S.A.R. e estudante de mestrado em Computação Aplicada na UFRPE. Como amante da Astronomia, sua pesquisa atual combina aprendizado de máquina e modelagem de ondas gravitacionais. Tem experiência, principalmente, com análise de séries temporais.



Emory R V Freitas

Emory Raphael Viana Freitas é engenheiro de software do Centro de Estudos e Sistemas Avançados do Recife (CESAR). Ele recebeu graduação em Ciência da Computação pela Universidade Federal do Amazonas em 2011, mestrado em Ciência da Computação pela Universidade Federal de Amazonas, em 2015.





[Titanic - Machine Learning from Disaster](#)

kaggle



Competir e competir!!!!



kaggle



Problema

[Ficheiro: Titanic survivors on the Carpathia, 1912.jpg - Wikipédia](#)



PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S
5	6	0	3	Moran, Mr. James	male	NaN	0	0	330877	8.4583	NaN	Q
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	E46	S



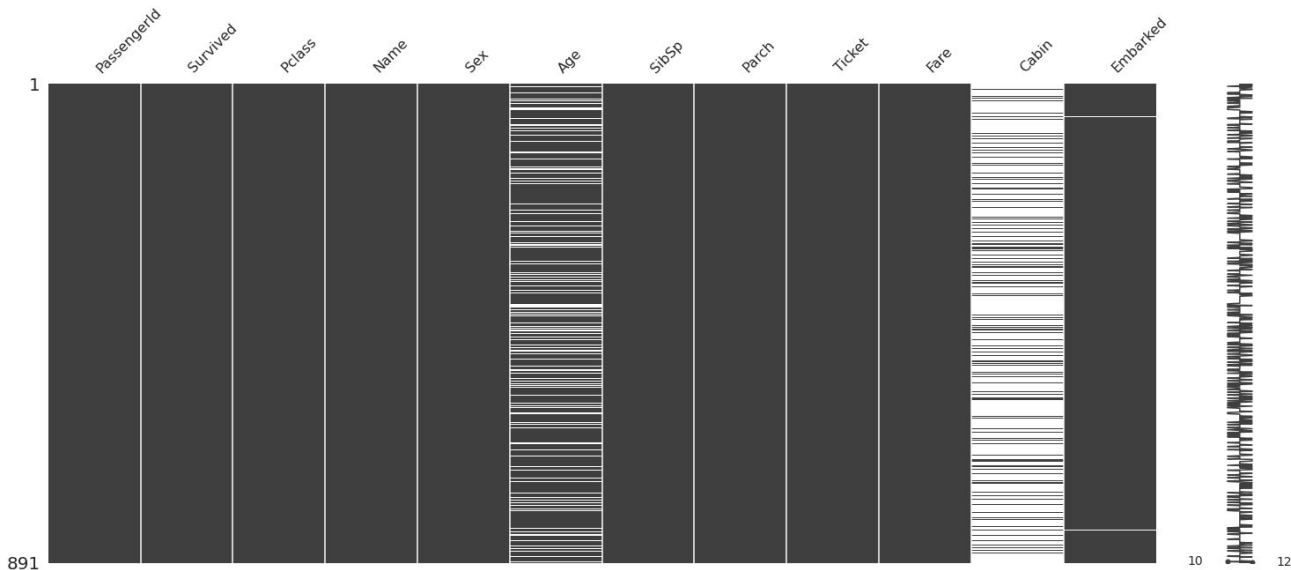
Análise e exploração de dados

Variable	Definition	Key
survival	Sobrevivência	0=No, 1=Sim
pclass	Classe de ingresso	1 = 1st, 2 = 2nd, 3 = 3rd
sex	Sexo	
Age	Idade em anos	
sibsp	# de irmãos/cônjuges	
parch	# de pais/filhos	
ticket	Número do bilhete	
fare	Tarifa de passageiro	
cabin	Número da cabine	
embarked	Porto de embarcação	C=Cherbourg, Q=Queenstown, S=Southampton



Análise e exploração de dados

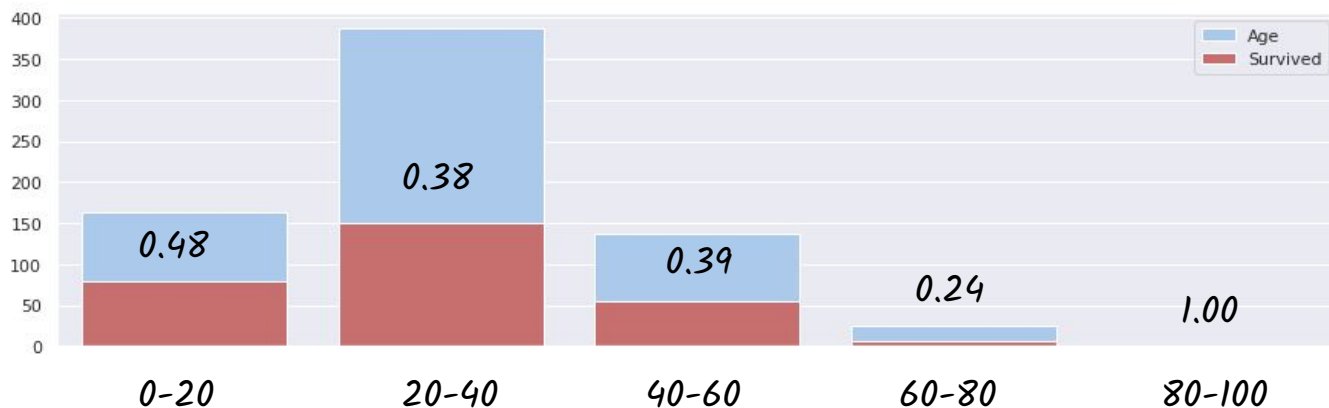
Missing value



Análise e exploração de dados

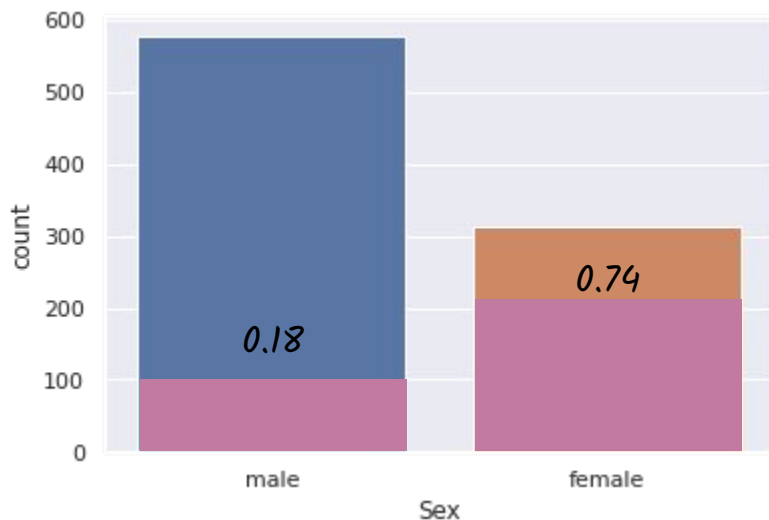
Análisis de Edad

$$x_{age} = \left\lfloor \frac{x_{age}}{q} \right\rfloor * q$$



Análise e exploração de dados

Análisis do sexo



Análise e exploração de dados

Seleção de features

	<i>X</i>						<i>y</i>
	Pclass	SibSp	Parch	Sex_female	Sex_male		
0	3	1	0	0	1		0
1	1	1	0	1	0		1
2	3	0	0	1	0		1
3	1	1	0	1	0		1
4	3	0	0	0	1		0
...
886	2	0	0	0	1		0
887	1	0	0	1	0		1
888	3	1	2	1	0		0
889	1	0	0	0	1		1
890	3	0	0	0	1		0



Treinamento do modelo

Dataset

Train

Test

891 examples

418 examples

<https://www.kaggle.com/c/titanic/leaderboard>



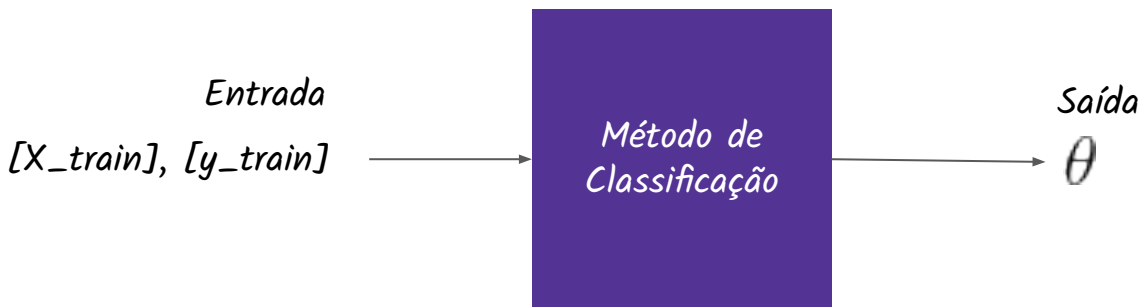
Treinamento do modelo

Dataset



Treinamento do modelo

Seleção do método de classificação

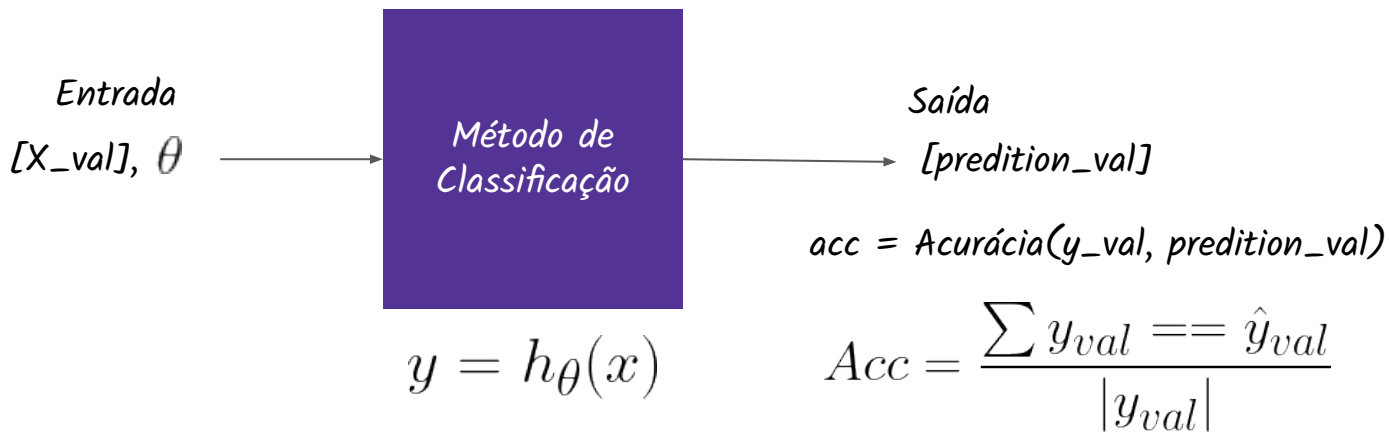


$$y = h_{\theta}(x)$$



Experimentos e Avaliação de Classificadores

Métrica acurácia



Objetivo

💡 *Em Machine Learning, um dos principais objetivos é estimar como um modelo irá se comportar em um ambiente real - i.e. como os seus resultados serão generalizados para um conjunto não visto de dados.*

Experimentos e Avaliação de Classificadores

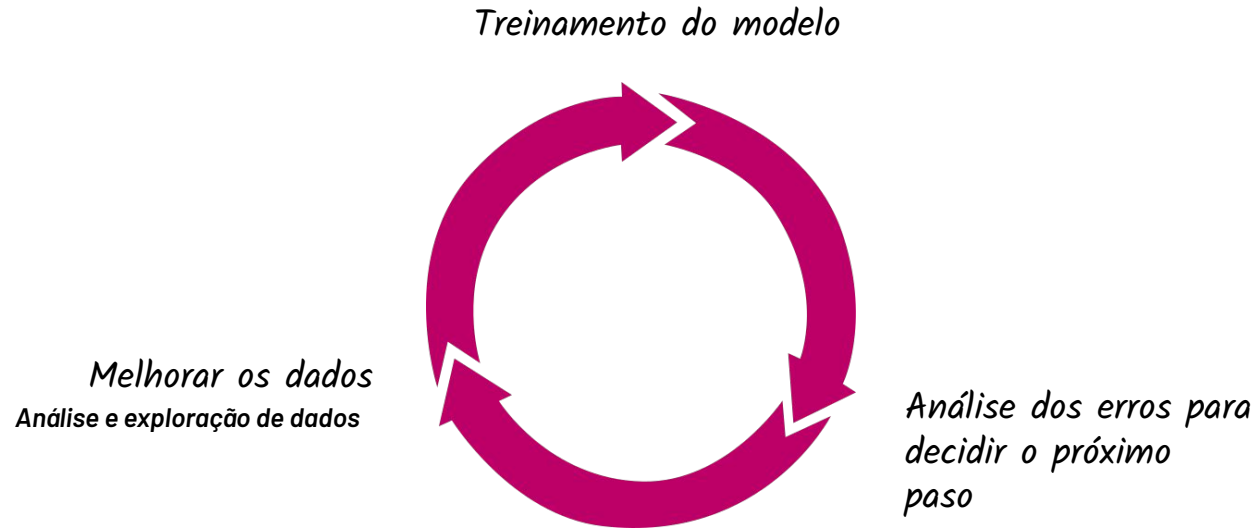
Mostrar nuevas métricas de avaliação de classificadores que permitam escolher de forma acertadas de modelos de classificação.

Competir e competir!!!!

As competições de Kaggle são uma forma muito boa de adquirir conhecimento na área de Machine Learning.



Workflow iterativo



Avaliação - Classificação

Na base do Titanic, temos como objetivo identificar os sobreviventes.

Survived?

<i>True</i>	<i>False</i>
-------------	--------------

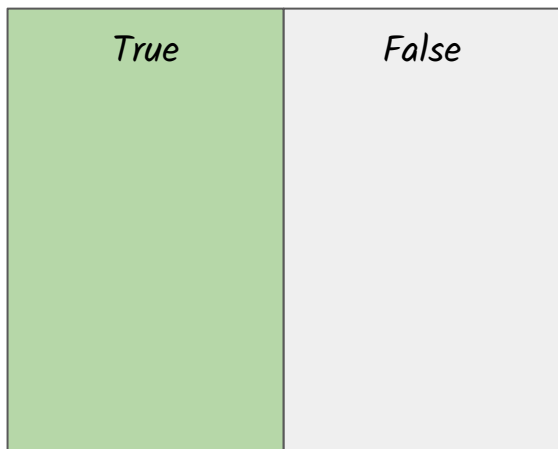
Conjunto verdade



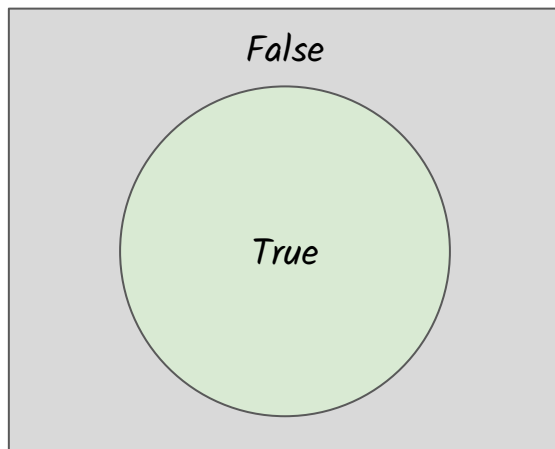
Avaliação - Classificação

Na base do Titanic, temos como objetivo identificar os sobreviventes.
O seu classificador identificou estes possíveis sobreviventes.

Survived?



Conjunto verdade

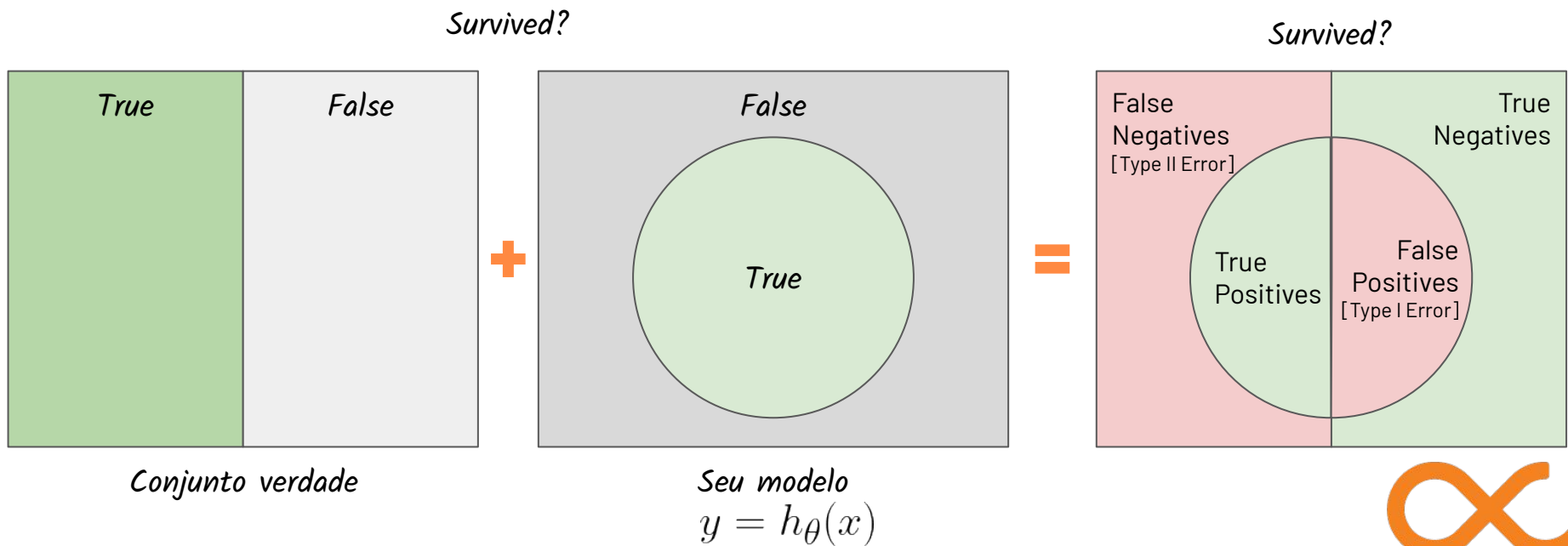


Seu modelo
 $y = h_{\theta}(x)$



Avaliação - Classificação

Como iremos avaliar este classificador?



Avaliação - Erros do tipo I e II

Type I Error

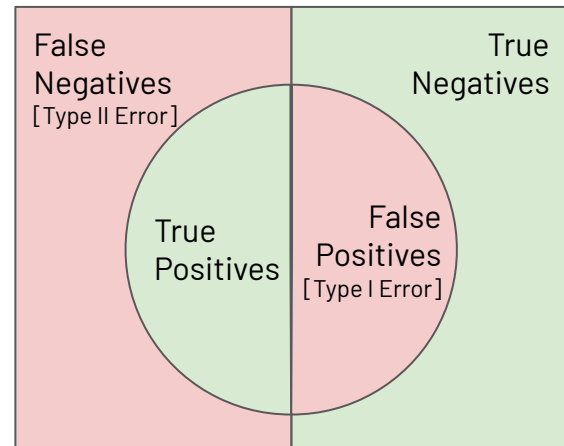


Type II Error



Avaliação – Matriz Confusão

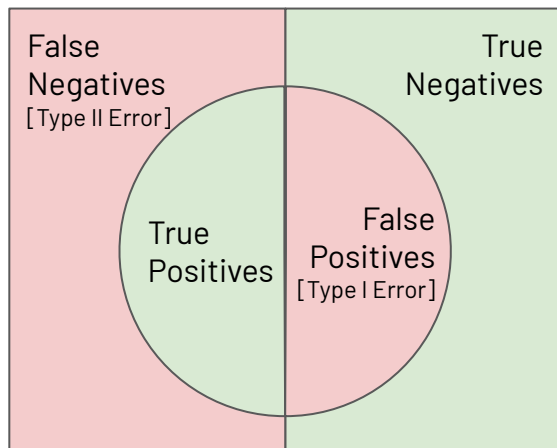
Os resultados serão avaliados através de uma **matriz confusão** e quantificados através de métricas diversas



	Identificado como sobrevivente	Identificado como não sobrevivente
Sobreviveu	Verdadeiro Positivo True Positive (TP)	Falso Negativo False Negative (FN) - E2
Não Sobreviveu	Falso Positivo False Positive (FP) - E1	Verdadeiro Negativo True Negative (TN)



Avaliação – Precisão e cobertura



$$\text{Precisão} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

Precision

Quando o modelo classificou um elemento como sendo da classe em questão, quanto ele acertou?

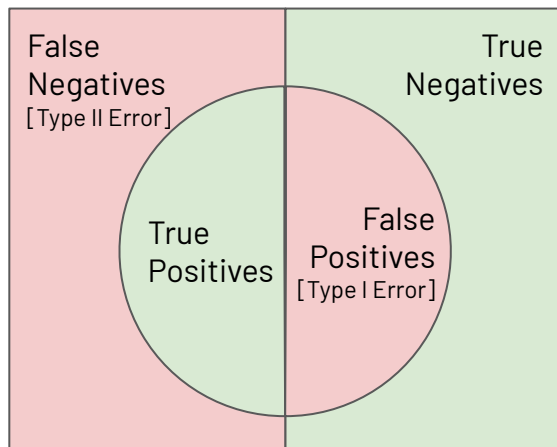
$$\text{Cobertura} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

Recall = Sensitivity

Da classe em questão, quantos itens o modelo acertou?



Avaliação - Acurácia



$$\text{Acurácia} = \frac{\text{Accuracy}}{\text{Accuracy}}$$

The diagram illustrates the formula for Accuracy. The numerator is a green circle partially overlapping a green square. The denominator is a green square with a circle inside it, where the left half of the circle is green and the right half is pink.


Quantos acertos fizemos no total?





Avaliação – Matriz Confusão

Os resultados serão avaliados através de uma **matriz confusão** e quantificados através das métricas vistas

	Identificado sobrevivente	Identificado não sobrevivente
Sobreviveu	TP	FN
Não Sobreviveu	FP	TN


$$\text{Cobertura} = TP / (FN + TP)$$


$$\text{Precisão} = TP / (TP + FP)$$


$$\text{Acurácia} = (TP + TN) / \text{Total}$$



Avaliação – Classificação

Como avaliar o classificador abaixo na prática do Titanic?

	Identificado como sobrevivente	Identificado como não sobrevivente
Sobreviveu	148	15
Não Sobreviveu	39	65



Avaliação – Classificação

Qual percentual de acerto do meu modelo para cada grupo?

Quando meu modelo afirmou algo, quão preciso ele foi?

	Identificado como sobrevivente	Identificado como não sobrevivente
Sobreviveu	148	15
Não Sobreviveu	39	65



Avaliação - Classificação

Qual percentual de acerto do meu modelo para cada grupo?

Quando meu modelo afirmou algo, quão preciso ele foi?

	Identificado como sobrevivente	Identificado como não sobrevivente
Sobreviveu	148	15
Não Sobreviveu	39	65

$$\text{Precisão (Classe 0)} = \frac{TP}{TP + FP}$$

$$\text{Precisão (Classe 1)} = \frac{TN}{TN + FN}$$



Avaliação - Classificação

Qual percentual de acerto do meu modelo para cada grupo?

Quando meu modelo afirmou algo, quão preciso ele foi?

	Identificado como sobrevivente	Identificado como não sobrevivente
Sobreviveu	148	15
Não Sobreviveu	39	65

$$\text{Precisão (Sobreviventes)} = 148 / (148 + 39) = 79\%$$

$$\text{Precisão (Não Sobreviventes)} = 65 / (65 + 15) = 81\%$$



Avaliação – Classificação

Qual o percentual de acerto para cada classe?

Olhando para cada classe isoladamente, qual a cobertura do nosso modelo?

	Identificado como sobrevivente	Identificado como não sobrevivente
Sobreviveu	148	15
Não Sobreviveu	39	65



Avaliação - Classificação

Qual o percentual de acerto para cada classe?

Olhando para cada classe isoladamente, qual a cobertura do nosso modelo?

	Identificado como sobrevivente	Identificado como não sobrevivente
Sobreviveu	148	15
Não Sobreviveu	39	65

Recall (Classe 0) ou Sensitivity =
 $TP / (FN + TP)$

Recall (Classe 1) ou Specificity =
 $TN / (FP + TN)$



Avaliação - Classificação

Qual o percentual de acerto para cada classe?

Olhando para cada classe isoladamente, qual a cobertura do nosso modelo?

	Identificado como sobrevivente	Identificado como não sobrevivente
Sobreviveu	148	15
Não Sobreviveu	39	65

$$\text{Recall (Classe 0) ou Sensitivity} = 148 / (15 + 148) = 91\%$$

$$\text{Recall (Classe 1) ou Specificity} = 65 / (39 + 65) = 62,5\%$$



Avaliação – Classificação

Qual percentual acertamos se considerarmos todas as classes do problema?

Quando meu modelo afirmou todas as classes, quão preciso ele foi?

	Identificado como sobrevivente	Identificado como não sobrevivente
Sobreviveu	148	15
Não Sobreviveu	39	65



Avaliação - Classificação

Qual percentual acertamos se considerarmos todas as classes do problema?

Quando meu modelo afirmou todas as classes, quão preciso ele foi?

	Identificado como sobrevivente	Identificado como não sobrevivente
Sobreviveu	148	15
Não Sobreviveu	39	65

$$\text{Accuracy} = (TP + TN) / \text{Total}$$



Avaliação - Classificação

Qual percentual acertamos se considerarmos todas as classes do problema?

Quando meu modelo afirmou todas as classes, quão preciso ele foi?

	Identificado como sobrevivente	Identificado como não sobrevivente
Sobreviveu	148	15
Não Sobreviveu	39	65

$$\text{Accuracy} = (148 + 65) / (148 + 65 + 39 + 15) = 80\%$$



Avaliação – Classificação

F1-Score: Média harmônica entre precisão e cobertura

	Identificado como sobrevivente	Identificado como não sobrevivente
Sobreviveu	148	15
Não Sobreviveu	39	65

$$F1\text{-Score} = \frac{2 * (Precision * Recall)}{(Precision + Recall)}$$



Avaliação – Classificação

F1-Score: Média harmônica entre precisão e cobertura

	Identificado como sobrevivente	Identificado como não sobrevivente
Sobreviveu	148	15
Não Sobreviveu	39	65

$$F1\text{-Score (Sobrevivente)} = 2 * (0,79 * 0,91) / (0,79 + 0,91) = 85\%$$

$$F1\text{-Score (Não Sobrevivente)} = 2 * (0,81 * 0,62) / (0,81 + 0,62) = 70\%$$



Avaliação – Matriz Confusão



Resumindo:

- **Precision** (Precisão): Número de itens corretamente identificados da classe em questão
- **Recall** / Sensitivity (Cobertura): Número de itens corretamente identificados considerando apenas a classe em questão
- **Accuracy** = (Acurácia) : Qual o percentual meu classificador acerta considerando todas as classes do problema?
- **F1-Score**: É uma média harmônica da precisão e recall



[Warning] Matriz Confusão

Na literatura ou em alguma biblioteca, a Matriz **Confusão** pode estar rotacionada!

		Actual Value (as confirmed by experiment)	
		positives	negatives
Predicted Value (predicted by the test)	positives	TP True Positive	FP False Positive
	negatives	FN False Negative	TN True Negative



[Warning] Matriz Confusão

		Actual Value (as confirmed by experiment)	
		positives	negatives
Predicted Value (predicted by the test)	positives	TP True Positive	FP False Positive
	negatives	FN False Negative	TN True Negative



$$\text{Precision} = TP / (TP + FP)$$



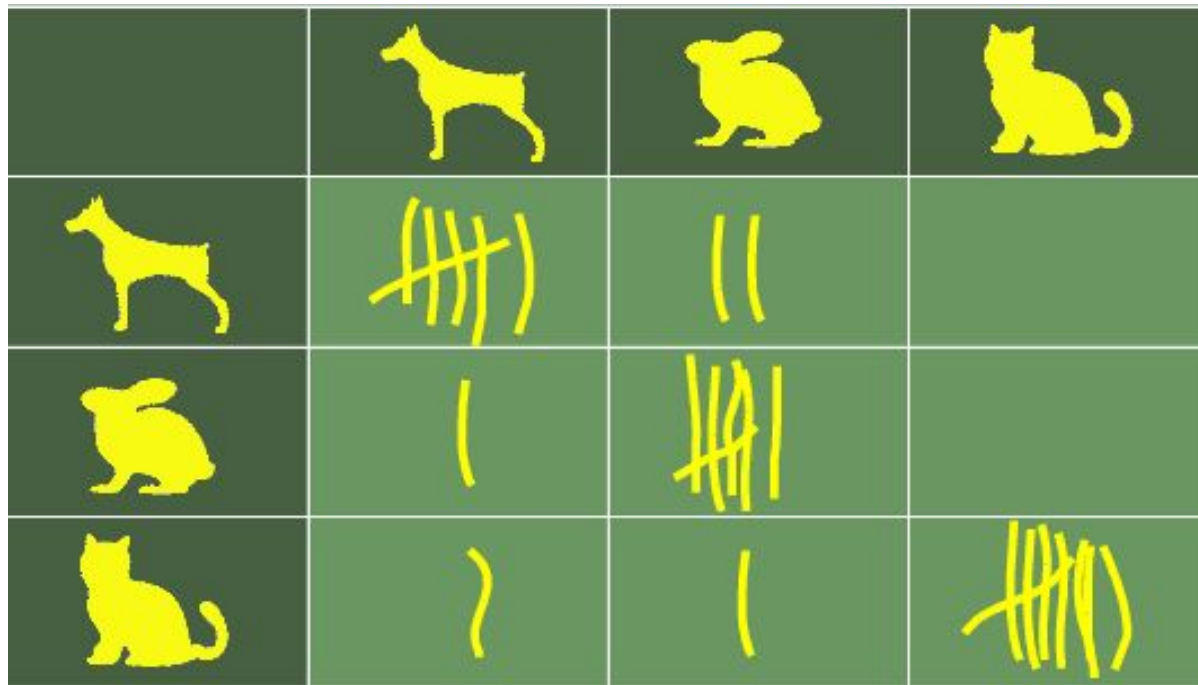
$$\text{Sensitivity/ Recall} = TP / (FN + TP)$$



$$\text{Accuracy} = (TP + TN) / \text{Total}$$



Avaliação - Classificação Multiclases



Avaliação

Verdadeiro

Predito

	C1	C2	C3
C1	14	2	0
C2	8	12	3
C3	0	0	19

C1

❖ Precision = ?

❖ Recall = ?



Avaliação

C1



Precision = $14 / 14 + 2 + 0 = 87\%$



Recall =

Verdadeiro

Predito	Verdadeiro			
		C1	C2	C3
	C1	14	2	0
	C2	8	12	3
	C3	0	0	19



Avaliação

C1

Verdadeiro

Predito

	C1	C2	C3
C1	14	2	0
C2	8	12	3
C3	0	0	19



Precision = $14 / 14 + 2 + 0 = 87\%$



Recall = $14 / 14 + 8 + 0 = 63\%$



Avaliação

Verdadeiro

Predito

	C1	C2	C3
C1	14	2	0
C2	8	12	3
C3	0	0	19

C1

- ❖ Precision = $14 / 14 + 2 + 0 = 87\%$
- ❖ Recall = $14 / 14 + 8 + 0 = 63\%$

Exercício: C2 e C3

- ❖ Precision = ?
- ❖ Recall = ?

Exercício: Geral

- ❖ Accuracy = ?



Avaliação

Verdadeiro

Predito		C1	C2	C3
	C1	14	2	0
	C2	8	12	3
	C3	0	0	19

C1

- ❖ Precision = $14 / 14 + 2 + 0 = 87\%$
- ❖ Recall = $14 / 14 + 8 + 0 = 63\%$

Exercício: C2 e C3

C2

- ❖ Precision = $12 / (12 + 8 + 3) = 52\%$
- ❖ Recall = $12 / (12 + 2) = 86\%$

C3

- ❖ Precision = $19 / 19 = 100\%$
- ❖ Recall = $19 / (19 + 3) = 86\%$



Avaliação

Verdadeiro

Predito		C1	C2	C3
	C1	14	2	0
	C2	8	12	3
	C3	0	0	19

Exercício: Geral

❖ Accuracy =

$$(14 + 12 + 19) / (14 + 2 + 8 + 12 + 3 + 19)$$

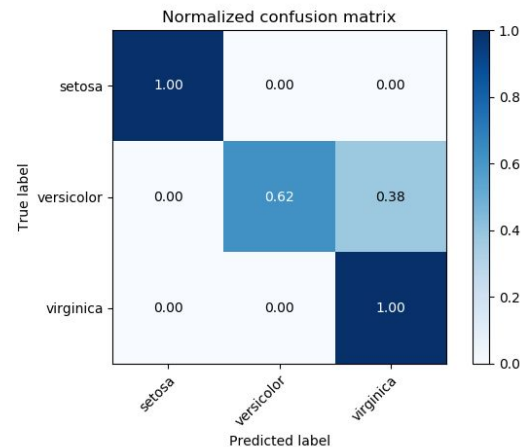
❖ Accuracy = 78%



[SKLearn] Matriz Confusão

```
from sklearn.metrics import confusion_matrix  
import seaborn as sns
```

```
cfm = confusion_matrix( y_test, y_pred )  
sns.heatmap(cfm, cbar=False, annot=True, cmap="Blues", fmt="d")
```



[SKLearn] Classification Report

```
from sklearn.metrics import classification_report  
print(classification_report( label, prediction ))
```

	precision	recall	f1-score	support
1	0.92	0.86	0.89	14
2	0.95	0.91	0.93	23
3	0.90	1.00	0.95	18
accuracy			0.93	55
macro avg	0.93	0.92	0.92	55
weighted avg	0.93	0.93	0.93	55

- macro average - averaging the unweighted mean per label
- weighted average - averaging the support-weighted mean per label





Hands On!

Experimentos em Machine
Learning com Python



Prática: Titanic - kNN

Vamos avaliar corretamente o kNN na base do Titanic.

1. Volte a prática do Titanic e avalie o seu modelo de acordo com as métricas abordadas nas bases de treino e teste: Matriz Confusão, Precisão, Cobertura, Acurácia e F1-Score
2. Compare e discuta os resultados
3. Descubra o melhor valor de k novamente, agora considerando outra métrica.



Para continuar...

- Katti Faceli et al. **Inteligência Artificial - Uma Abordagem de Aprendizado de Máquina.** LTC, 2019.
 - Capítulo 9. Avaliação de Modelos Preditivos:
9.1(.1), 9.2(.1 e .2), 9.3(.1)
- Aurélien Géron. **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems.** 2a Edição: O'Reilly Media, 2019.
 - Chapter 3. Classification
Performance Measures (Exceto Curva RoC)



Outros:

- [A Gentle Introduction to k-fold Cross-Validation](#)
- [Cross Validation: Avaliando seu modelo de Machine Learning](#) by Eduardo Braz Rabello
- [Pre-Process Data with Pipeline to Prevent Data Leakage during Cross-Validation](#) by Kai Zhao in @TDataScience
- [What is a Confusion Matrix in Machine Learning](#)
- [Understanding Confusion Matrix by Sarang Narkhede in @TDataScience](#)





C . E . S . A . R

Pessoas impulsionando inovação.
Inovação impulsionando negócios.

NOSSO CONTATO

cesar.org.br

cesar.school





C . E . S . A . R

sch∞l

