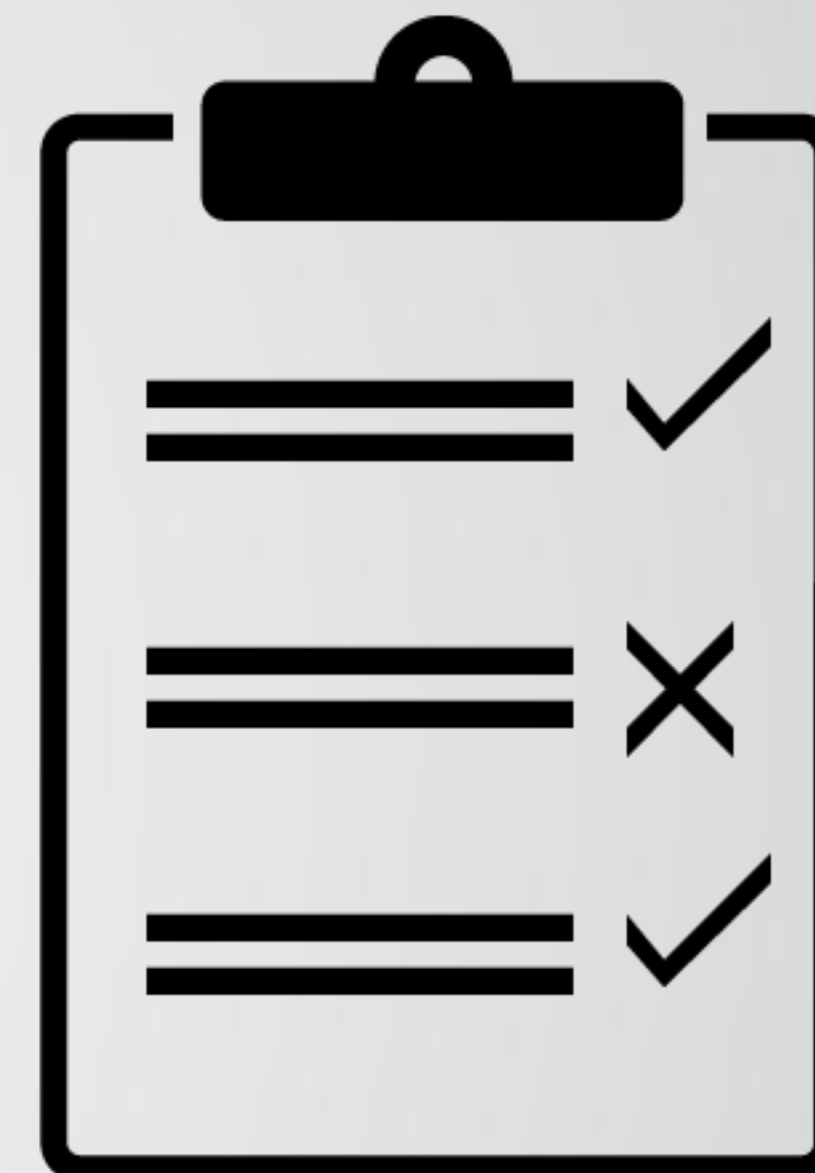


Sdco



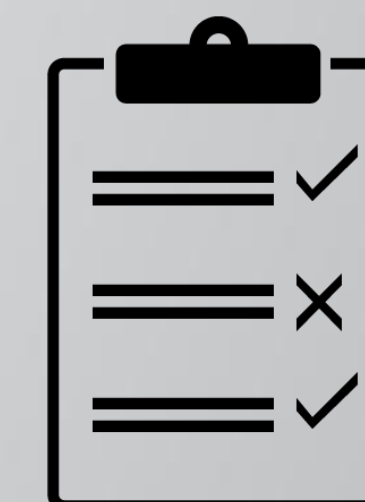
# MEDIDAS DE DESEMPENHO



# MEDIDAS DE DESEMPENHO



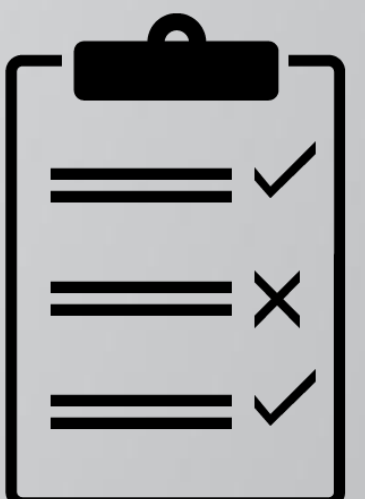
UM ALGORITMO É BOM SE PRODUZ  
HIPÓTESES QUE FAZEM UM BOM  
TRABALHO DE PREVISÃO DE  
CLASSIFICAÇÃO DE **EXEMPLOS NÃO-  
VISTOS**



# MEDIDAS DE DESEMPENHO



MACHINE LEARNING UMA FERRAMENTA  
PODEROSA PARA AQUISIÇÃO  
AUTOMÁTICA DE CONHECIMENTO,  
ENTRETANTO, **NÃO EXISTE UM ÚNICO**  
**ALGORITMO** QUE APRESENTE MELHOR  
DESEMPENHO PARA TODOS PROBLEMAS.

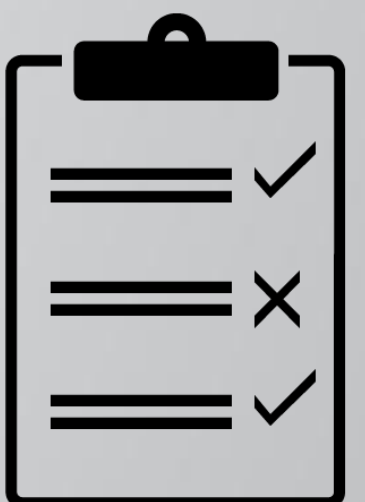


# COMO TESTAR UM ALGORITMO

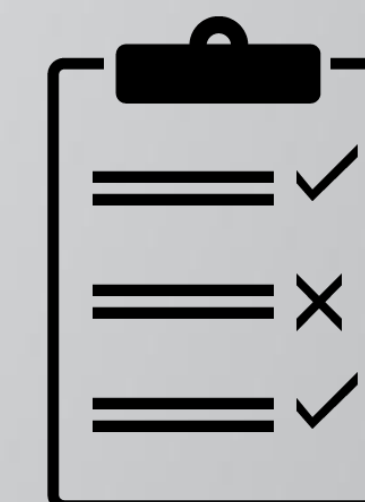


# PASSOS PARA **TESTE DESEMPENHO** Sd $\infty$

- 1 – COLETAR UM GRANDE CONJUNTO DE EXEMPLOS;
- 2 – DIVIDI-LO EM DOIS CONJUNTOS (TREINO/TESTE)
- 3 – TREINAR O ALGORITMO COM O CONJUNTO DE TREINO;
- 4 – TESTAR A HIPÓTESE COM O CONJUNTO DE TESTE;
- 5 – MEDIR OS ACERTOS DA HIPÓTESE.



# ERRO



## EQUAÇÃO

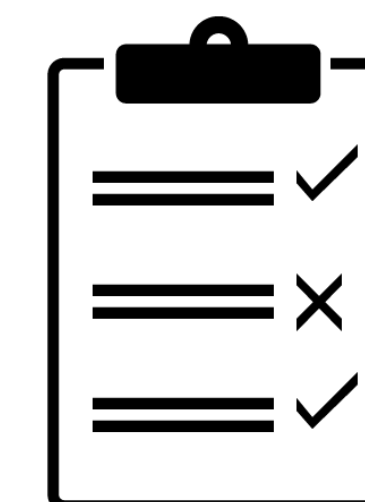
$$err(h) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{y_i \neq h(x_i)}$$

### EXEMPLO:

Em um conjunto com 100 objetos de teste 95 foram classificados corretamente e 5 não.

Então o erro será de

$5/100 = 0.05$  ou 5%





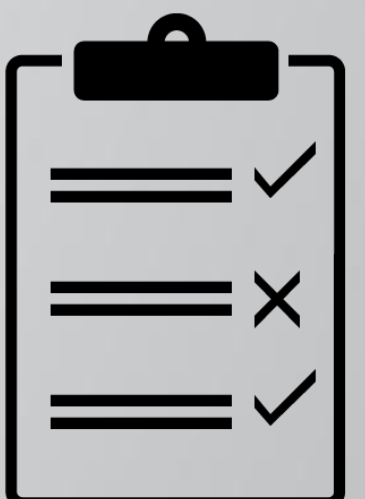
# ACURÁCIA



# ACURÁCIA



ENTRETANTO EM MACHINE LEARNING  
NÃO ENFATIZAMOS OS ERROS DO  
ALGORITMOS E SIM A TAXA DE ACERTOS  
COMO A ACURÁCIA.

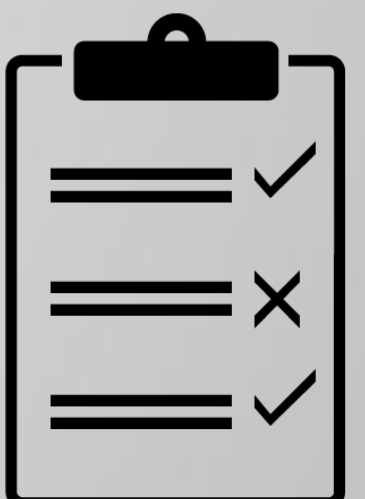


# ACURÁCIA



## DEFINIÇÃO:

PROXIMIDADE ENTRE O VALOR OBTIDO  
EXPERIMENTALMENTE E O VALOR VERDADEIRO  
NA MEDIÇÃO DE UMA GRANDEZA FÍSICA.



## EQUAÇÃO

$$acc(h) = 1 - err(h)$$

### EXEMPLO:

Em um conjunto com 100 objetos de teste 95 foram classificados corretamente e 5 não.

Então o erro será de  $5/100 = 0.05$

e a acurácia de  $1 - 0.05 = 0.95$

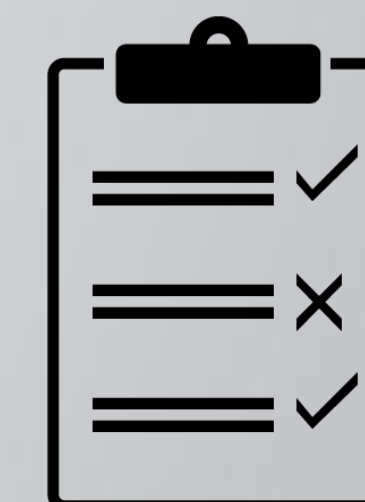
ou 95%

# VALIDAÇÃO CRUZADA

# VALIDAÇÃO CRUZADA



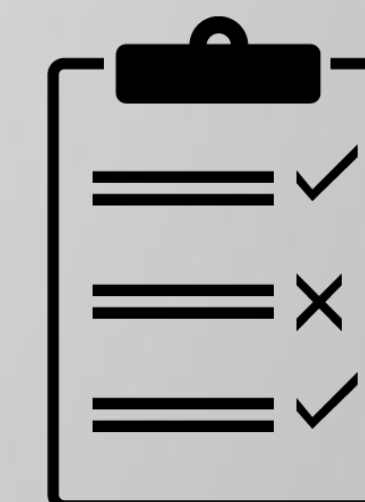
QUANDO DIVIDIMOS O CONJUNTO DE DADOS EM DOIS CONJUNTOS, PODEMOS REDUZIR DRASTICAMENTE O NÚMERO DE AMOSTRAS PARA O TREINAMENTO.



# VALIDAÇÃO CRUZADA



O CONCEITO CENTRAL DAS TÉCNICAS DE  
VALIDAÇÃO CRUZADA É O  
PARTICIONAMENTO DO CONJUNTO DE  
DADOS EM SUBCONJUNTOS  
MUTUALMENTE EXCLUSIVOS.



# VALIDAÇÃO CRUZADA

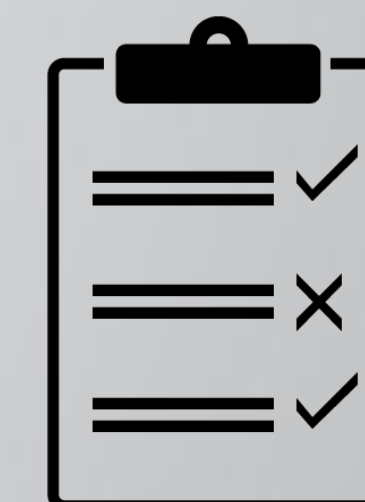


**BASICAMENTE 3 TIPOS:**

*HOLDOUT;*

*K-FOLD;*

*LEAVE-ONE-OUT.*





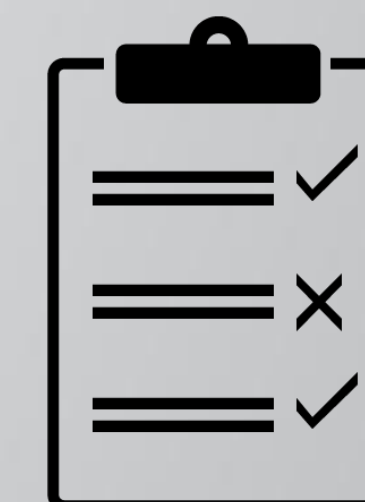
# VALIDAÇÃO CRUZADA



## HOLDOUT

Dividir o conjunto de dados em dois.

Muito comum é considerar  $\frac{2}{3}$  dos dados para treinamento e o  $\frac{1}{3}$  restante para teste.



## SCIKIT LEARN

IMPORTAR **TRAIN\_TEST\_SPLIT** PARA DIVIDIR  
O CONJUNTO DE DADOS EM TREINO E TESTE

```
from sklearn.model_selection import train_test_split
```

## SCIKIT LEARN

### EXECUTAR A DIVISÃO DOS CONJUNTOS

```
X_train, X_test, y_train, y_test=train_test_split(X, y, test_size=0.4)
```

X\_train = Conjunto de treino das Features

X\_test = Conjunto de teste das Features

y\_train = Conjunto de treino das targets

y\_test = Conjunto de teste das targets

## SCIKIT LEARN

### EXECUTAR A DIVISÃO DOS CONJUNTOS

```
X_train, X_test, y_train, y_test=train_test_split(X,y,test_size=0.4)
```

X = Conjunto original das Features

Y = Conjunto original das targets

test\_size = porcentagem do conjunto de teste

## SCIKIT LEARN

### Executando KNN

```
from sklearn.neighbors import KNeighborsClassifier
```

```
knn = KNeighborsClassifier(n_neighbors=3)
```

```
knn.fit(X_train, y_train)
```

```
knn.score(X_test, y_test)
```

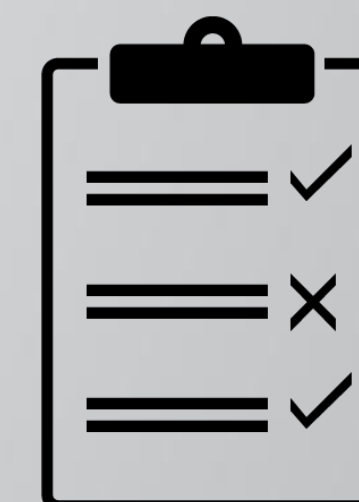
## SCIKIT LEARN

Executando com árvores de decisão

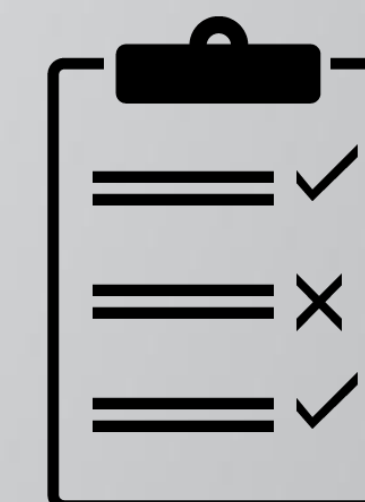
```
from sklearn import tree  
  
tree = tree.DecisionTreeClassifier()  
  
tree.fit(X_train, y_train)  
  
tree.score(X_test, y_test)
```

UMA

# MELHOR DIVISÃO



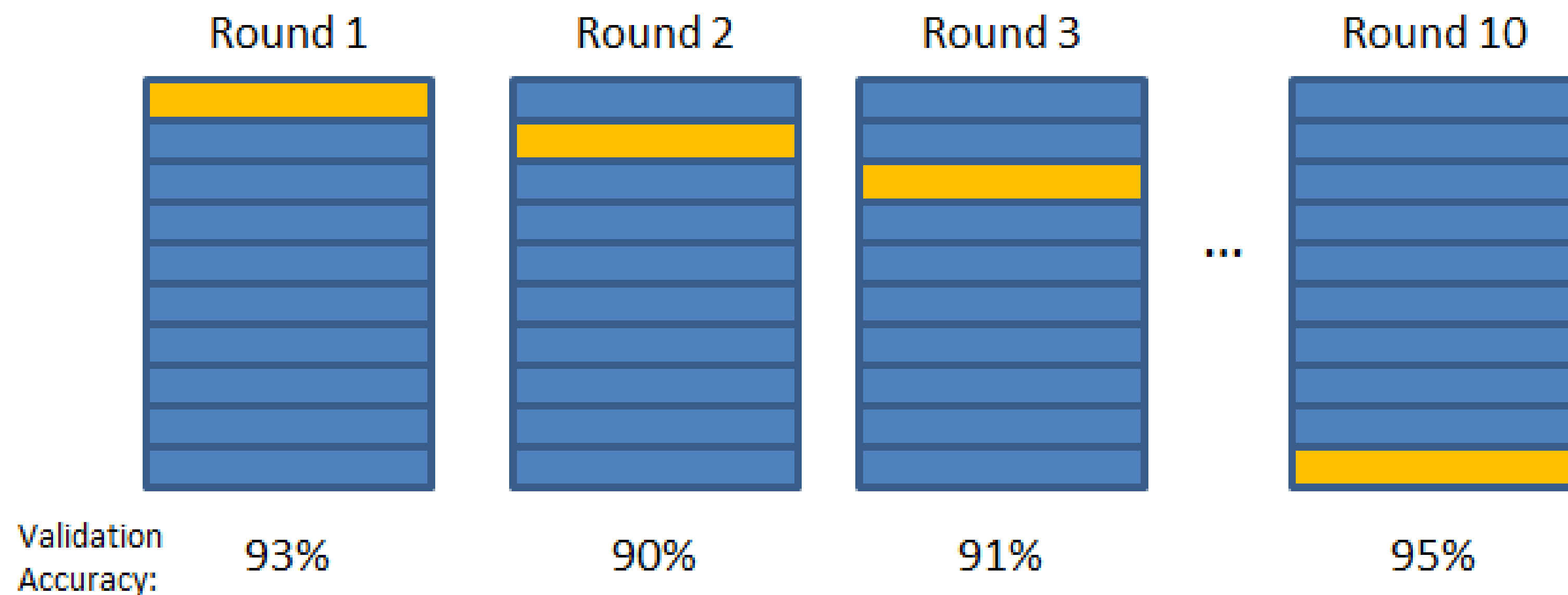
DIVIDIR O CONJUNTO TOTAL DE DADOS  
EM **K** SUBCONJUNTOS MUTUAMENTE  
EXCLUSIVOS DO MESMO TAMANHO.





# K-FOLD

Validation Set  
Training Set



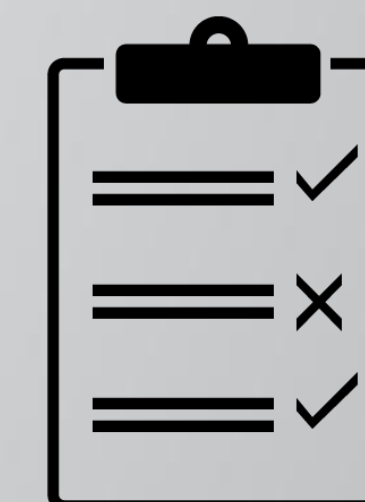
Final Accuracy = Average(Round 1, Round 2, ...)

Partir disto, um subconjunto é utilizado para teste e os k-1 para treino. Calcula-se a acurácia do modelo. Este processo é realizado k vezes alternando de forma circular o subconjunto de teste.

# LEAVE-ONE-OUT



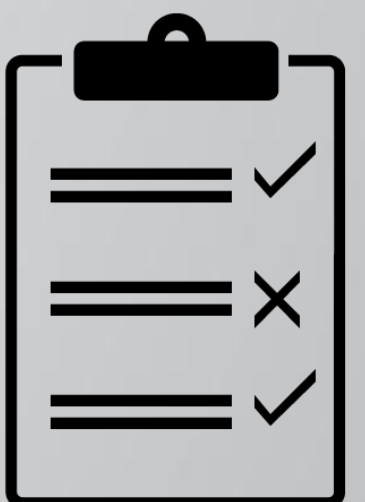
É UM CASO ESPECÍFICO DO **K-FOLD**,  
COM **K** IGUAL AO NÚMERO TOTAL DE  
DADOS **N**. NESTA ABORDAGEM SÃO  
REALIZADOS **N** CÁLCULOS DE ERRO, UM  
PARA CADA DADO.



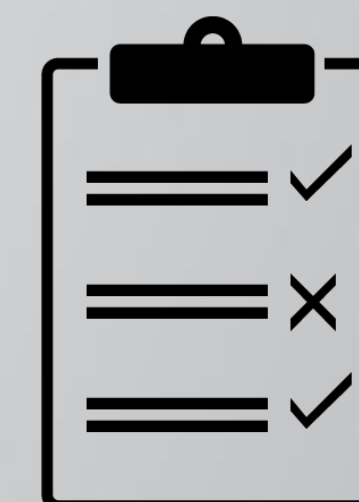
# LEAVE-ONE-OUT



APESAR DA INVESTIGAÇÃO COMPLETA  
DO MODELO EM RELAÇÃO AOS DADOS  
UTILIZADOS, POSSUI ALTO CUSTO  
COMPUTACIONAL.



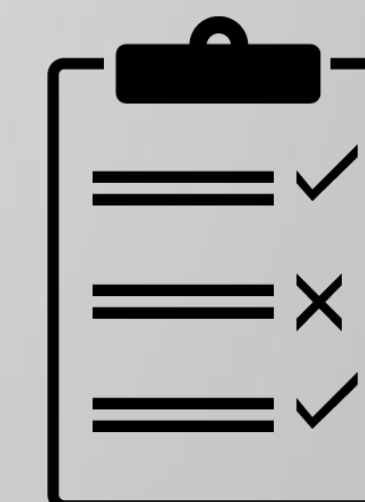
# MATRIZ DE CONFUSÃO



# MATRIZ DE CONFUSÃO



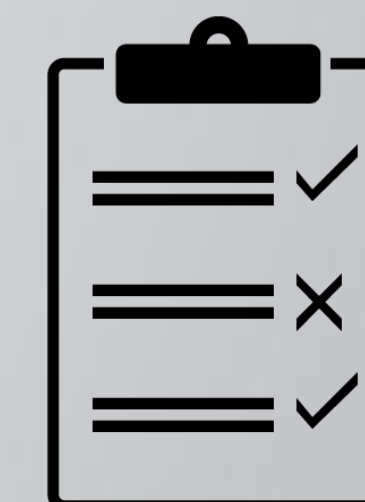
A **ACURÁCIA** NÃO É UMA MÉTRICA  
CONFIÁVEL PARA O DESEMPENHO REAL  
DE UM CLASSIFICADOR.



# MATRIZ DE CONFUSÃO

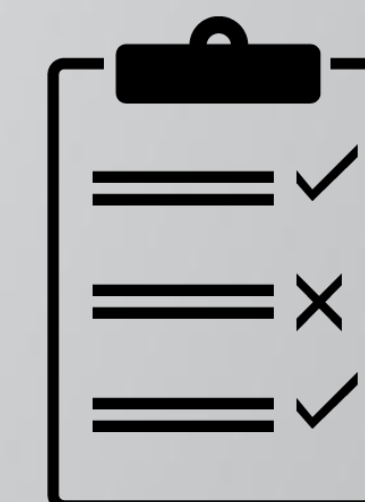
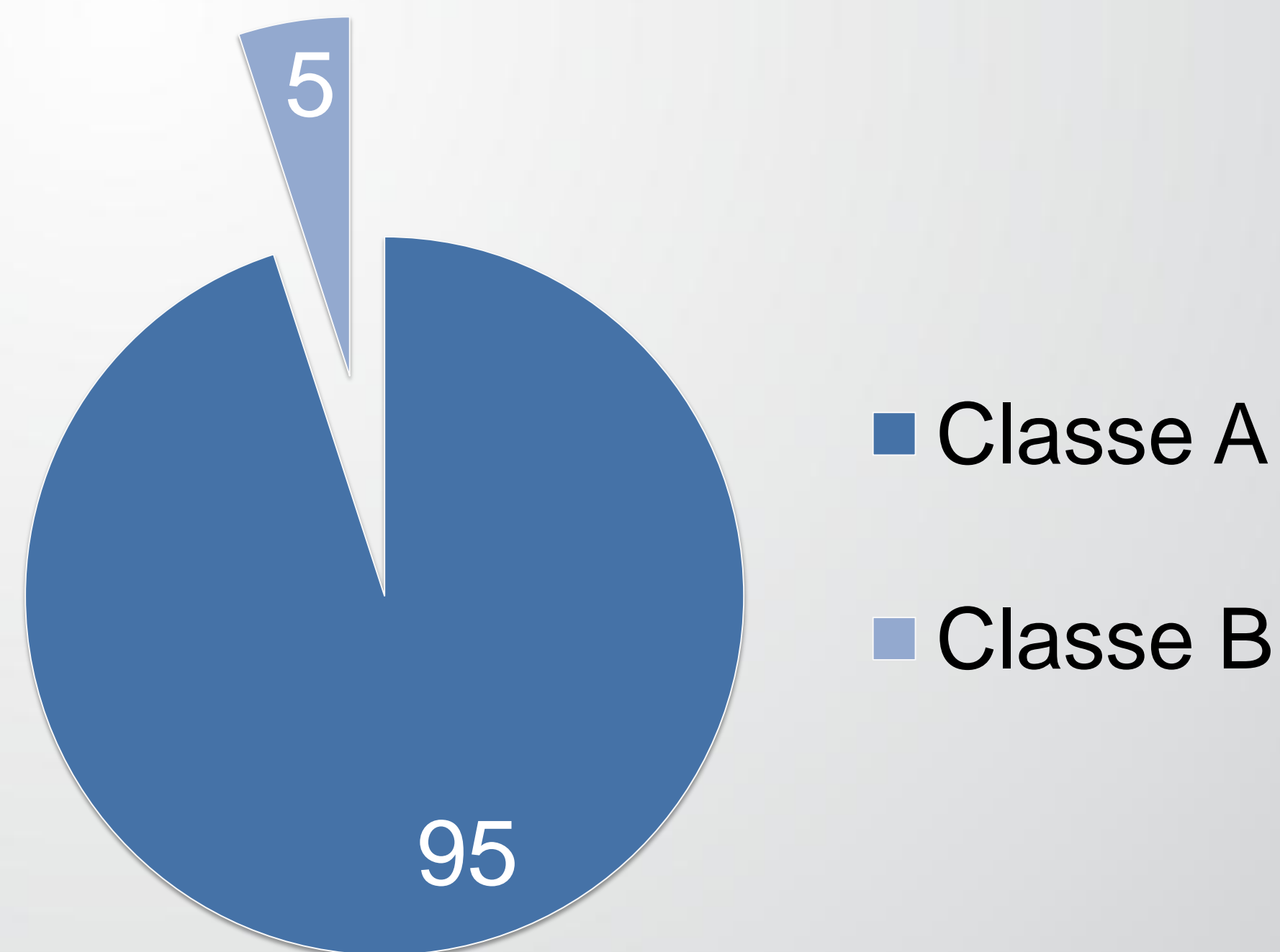


PORQUE PRODUZIRÁ RESULTADOS  
ENGANADORES SE O CONJUNTO DE  
DADOS ESTIVER DESEQUILIBRADO.



# MATRIZ DE CONFUSÃO

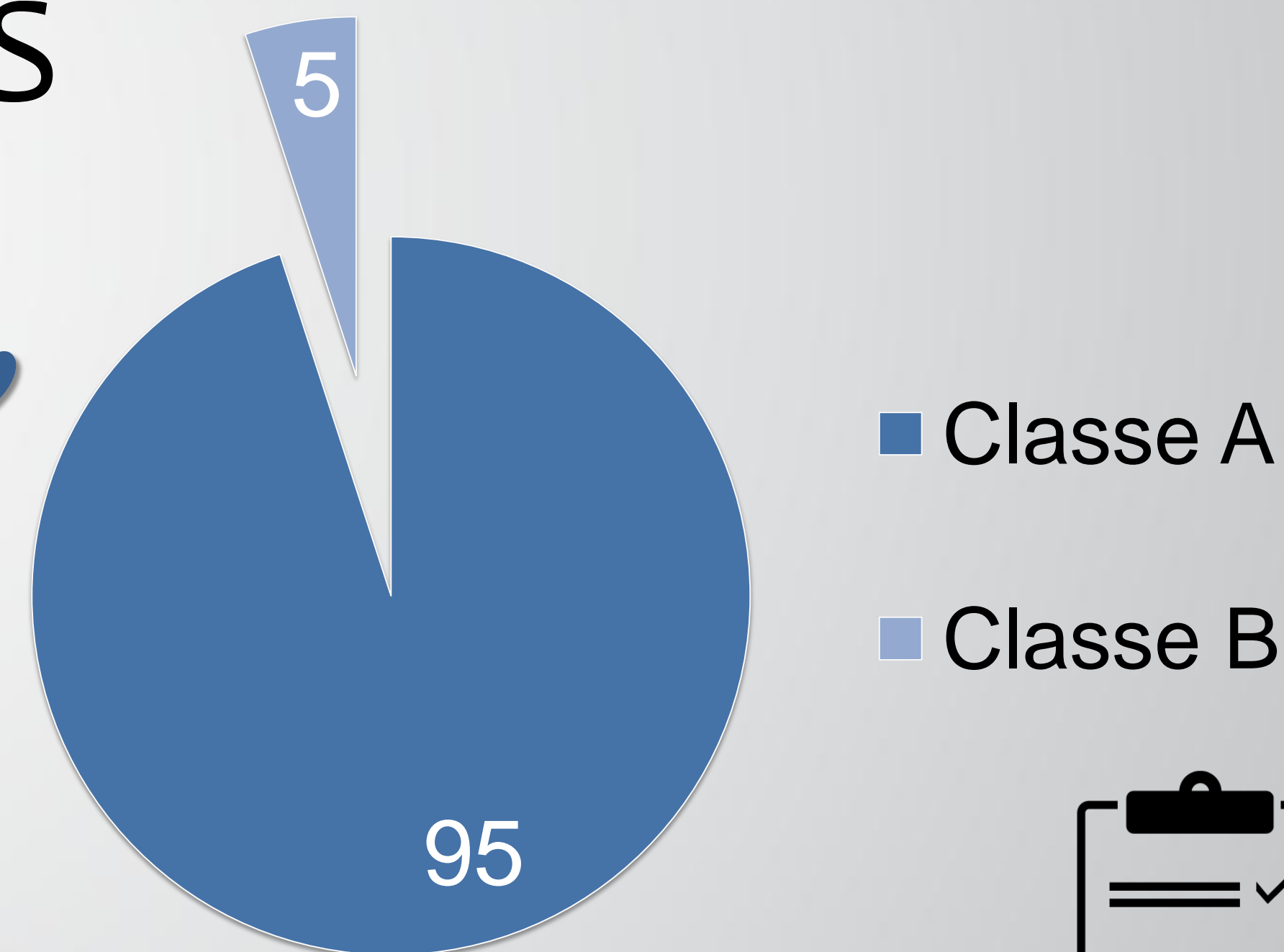
POR EXEMPLO, SE HOUVESSE 95  
EXEMPLOS DE UMA **CLASSE A** E APENAS  
5 DE OUTRA **CLASSE B**





# MATRIZ DE CONFUSÃO

UM CLASSIFICADOR PARTICULAR  
PODERIA CLASSIFICAR TODAS AS  
OBSERVAÇÕES COMO **CLASSE A**,  
DADO O DESEQUILÍBRIO

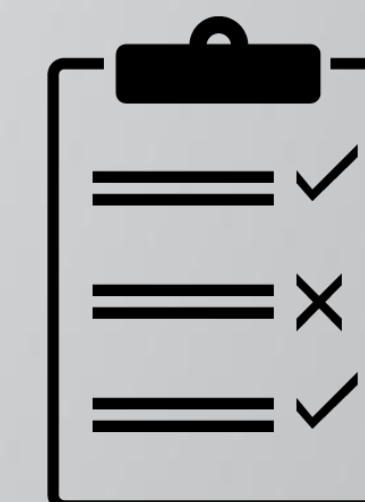




# MATRIZ DE CONFUSÃO



DADO UM EXPERIMENTO  
DE INSTÂNCIAS **P** POSITIVAS E **N**  
INSTÂNCIAS NEGATIVAS A MATRIZ DE  
CONFUSÃO SERÁ:



# MATRIZ DE CONFUSÃO

		Condição verdadeira	
População total		Condição positiva	Condição negativa
Condição prevista	Condição preditiva positiva	<b>Verdadeiro positivo</b> VP	<b>Falso positivo</b> (erro Tipo I) FP
	Condição preditiva negativa	<b>Falso negativo</b> (erro Tipo II) FN	<b>Verdadeiro negativo</b> VN

# MATRIZ DE CONFUSÃO



## SCIKIT LEARN

Executando com KNN

```
from sklearn.model_selection import cross_val_predict
```

```
from sklearn.metrics import confusion_matrix
```

```
y_pred = cross_val_predict(nayB,X,y,cv=5)
```

```
conf_mat = confusion_matrix(y,y_pred)
```

# MATRIZ DE CONFUSÃO



MÉTRICAS GERADAS

A PARTIR DA MATRIZ DE CONFUSÃO



# MATRIZ DE CONFUSÃO



## ACURÁCIA

A PROPORÇÃO DE PREDIÇÕES CORRETAS, SEM LEVAR EM CONSIDERAÇÃO O QUE É POSITIVO E O QUE É NEGATIVO. ESTA MEDIDA É ALTAMENTE SUSCETÍVEL A DESBALANCEAMENTOS DO CONJUNTO DE DADOS E PODE FACILMENTE INDUZIR A UMA CONCLUSÃO ERRADA SOBRE O DESEMPENHO DO SISTEMA.

		Condição verdadeira	
		Condição positiva	Condição negativa
Condição prevista	Condição preditiva positiva	VP	FP
	Condição preditiva negativa	FN	VN

## ACURACIA

TOTAL DE ACERTOS /N DE EXEMPLOS  
 $(VP + VN) / (P + N)$

# MATRIZ DE CONFUSÃO



## ***SENSIBILIDADE ou REVOCAÇÃO***

A PROPORÇÃO DE VERDADEIROS POSITIVOS:

A CAPACIDADE DO SISTEMA EM PREDIZER

CORRETAMENTE A CONDIÇÃO PARA CASOS QUE

REALMENTE A TÊM.

		Condição verdadeira	
		Condição positiva	Condição negativa
Condição prevista	Condição preditiva positiva	VP	FP
	Condição preditiva negativa	FN	VN

## ***SENSIBILIDADE ou REVOCAÇÃO***

ACERTOS POSITIVOS / TOTAL DE POSITIVOS

$$VP / (VP + FN)$$

# MATRIZ DE CONFUSÃO



## ***ESPECIFICIDADE***

A PROPORÇÃO DE VERDADEIROS NEGATIVOS.

A CAPACIDADE DO SISTEMA EM PREDIZER  
CORRETAMENTE A AUSÊNCIA DA CONDIÇÃO  
PARA CASOS QUE REALMENTE NÃO A TÊM.

		Condição verdadeira	
		Condição positiva	Condição negativa
Condição prevista	Condição preditiva positiva	VP	FP
	Condição preditiva negativa	FN	VN

## ***ESPECIFICIDADE***

ACERTOS NEGATIVOS / TOTAL DE NEGATIVOS

$$VN / (VN + FP)$$



# MATRIZ DE CONFUSÃO



## **EFICIÊNCIA**

A MÉDIA ARITMÉTICA DA SENSIBILIDADE E ESPECIFICIDADE. NA PRÁTICA, A SENSIBILIDADE E A ESPECIFICIDADE VARIAM EM DIREÇÕES OPOSTAS. ISTO É, GERALMENTE, QUANDO UM MÉTODO É MUITO SENSÍVEL A POSITIVOS, TENDE A GERAR MUITOS FALSO-POSITIVOS, E VICE-VERSA. ASSIM, UM MÉTODO DE DECISÃO PERFEITO (100 % DE SENSIBILIDADE E 100% ESPECIFICIDADE) RARAMENTE É ALCANÇADO.

		Condição verdadeira	
		Condição positiva	Condição negativa
Condição prevista	População total		
	Condição preditiva positiva	VP	FP
	Condição preditiva negativa	FN	VN

## **EFICIÊNCIA**

$$(SENS + ESPEC) / 2$$



# MATRIZ DE CONFUSÃO



## **VALOR PREDITIVO POSITIVO ou PRECISÃO**

A PROPORÇÃO DE VERDADEIROS POSITIVOS EM  
RELAÇÃO A TODAS AS PREDIÇÕES POSITIVAS.

ESTA MEDIDA É ALTAMENTE SUSCETÍVEL A  
DESBALANCEAMENTOS DO CONJUNTO DE DADOS E  
PODE FACILMENTE INDUZIR A UMA CONCLUSÃO  
ERRADA SOBRE O DESEMPENHO DO SISTEMA.

		Condição verdadeira	
		Condição positiva	Condição negativa
Condição prevista	Condição preditiva positiva	VP	FP
	Condição preditiva negativa	FN	VN

## **VALOR PREDITIVO POSITIVO ou PRECISÃO**

ACERTOS POSITIVOS / TOTAL DE PREDIÇÕES POSITIVAS  
 $VP / (VP + FP)$

# MATRIZ DE CONFUSÃO



## **VALOR PREDITIVO NEGATIVO**

A PROPORÇÃO DE VERDADEIROS NEGATIVAS EM  
RELAÇÃO A TODAS AS PREDIÇÕES NEGATIVAS.

ESTA MEDIDA É ALTAMENTE SUSCETÍVEL A  
DESBALANCEAMENTOS DO CONJUNTO DE DADOS E  
PODE FACILMENTE INDUZIR A UMA CONCLUSÃO  
ERRADA SOBRE O DESEMPENHO DO SISTEMA.

		Condição verdadeira	
		Condição positiva	Condição negativa
Condição prevista	População total		
	Condição preditiva positiva	VP	FP
	Condição preditiva negativa	FN	VN

## **VALOR PREDITIVO NEGATIVO**

ACERTOS NEGATIVOS / TOTAL DE PREDIÇÕES NEGATIVAS  
 $VN / (VN + FN)$

# MATRIZ DE CONFUSÃO



## COEFICIENTE DE CORRELAÇÃO DE MATTHEWS – COEFICIENTE (PHI)

O coeficiente de correlação de Matthews é uma medida de qualidade de duas classificações binárias que pode ser usada mesmo se as classes possuem tamanhos bastante diferentes. Retorna um valor entre (-1) e (+1), em que um coeficiente de (+1) representa uma predição perfeita, (0) representa uma predição aleatória media, e (-1) uma predição inversa. Esta estatística é equivalente ao coeficiente phi, e tenta, assim como a eficiência, resumir a qualidade da tabela de contingência em um único valor numérico passível de ser comparado.

		Condição verdadeira	
		Condição positiva	Condição negativa
Condição prevista	Condição preditiva positiva	VP	FP
	Condição preditiva negativa	FN	VN

## VALOR PREDITIVO POSITIVO

$$(VP * VN - FP * FN) / \sqrt{((VP + FP) * (VP + FN) * (VN + FP) * (VN + FN))}$$

# MATRIZ DE CONFUSÃO



## ***F-MEASURE***

A MEDIDA QUE COMBINA PRECISÃO E REVOCAÇÃO É A MÉDIA HARMÓNICA DE PRECISÃO E REVOCAÇÃO, A TRADICIONAL F-MEASURE OU F-SCORE BALANCEADA

		Condição verdadeira	
		Condição positiva	Condição negativa
Condição prevista	População total		
	Condição preditiva positiva	VP	FP
Condição preditiva negativa		FN	VN

## ***F-MEASURE***

$$2 * ((Precisao * revocação) / (Precisao + revocação))$$

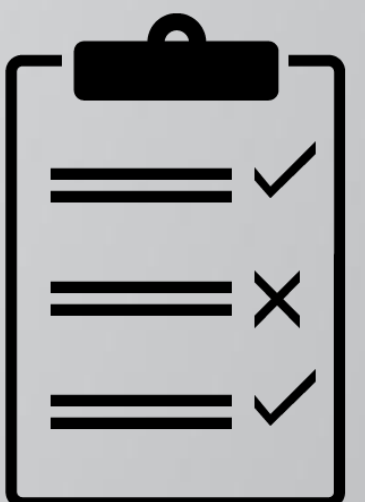
UMA

# METRICA VISUAL





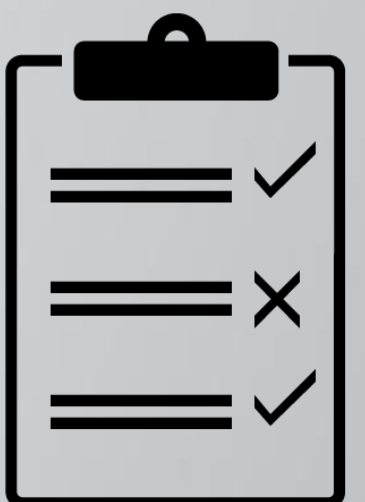
A CURVA ROC É UMA FERRAMENTA  
PODEROSA PARA MEDIR E ESPECIFICAR  
PROBLEMAS DE DESEMPENHO DOS  
CLASSIFICADORES.



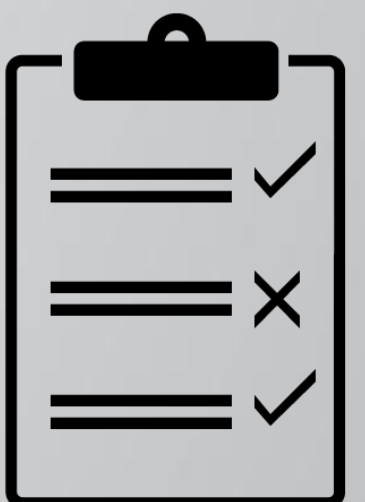
# CURVA ROC



É BASEADA EM MÉTRICAS DA MATRIZ  
DE CONFUSÃO COM SENSIBILIDADE E  
ESPECIFICIDADE.



QUANTO MAIOR A SENSIBILIDADE DO  
ALGORITMO DE CLASSIFICAÇÃO AO  
PROBLEMA MELHOR





# CURVA ROC

## EXEMPLO

A – PONTO IDEAL

B – BOA CLASSIFICAÇÃO

C – RUIM CLASSIFICAÇÃO

