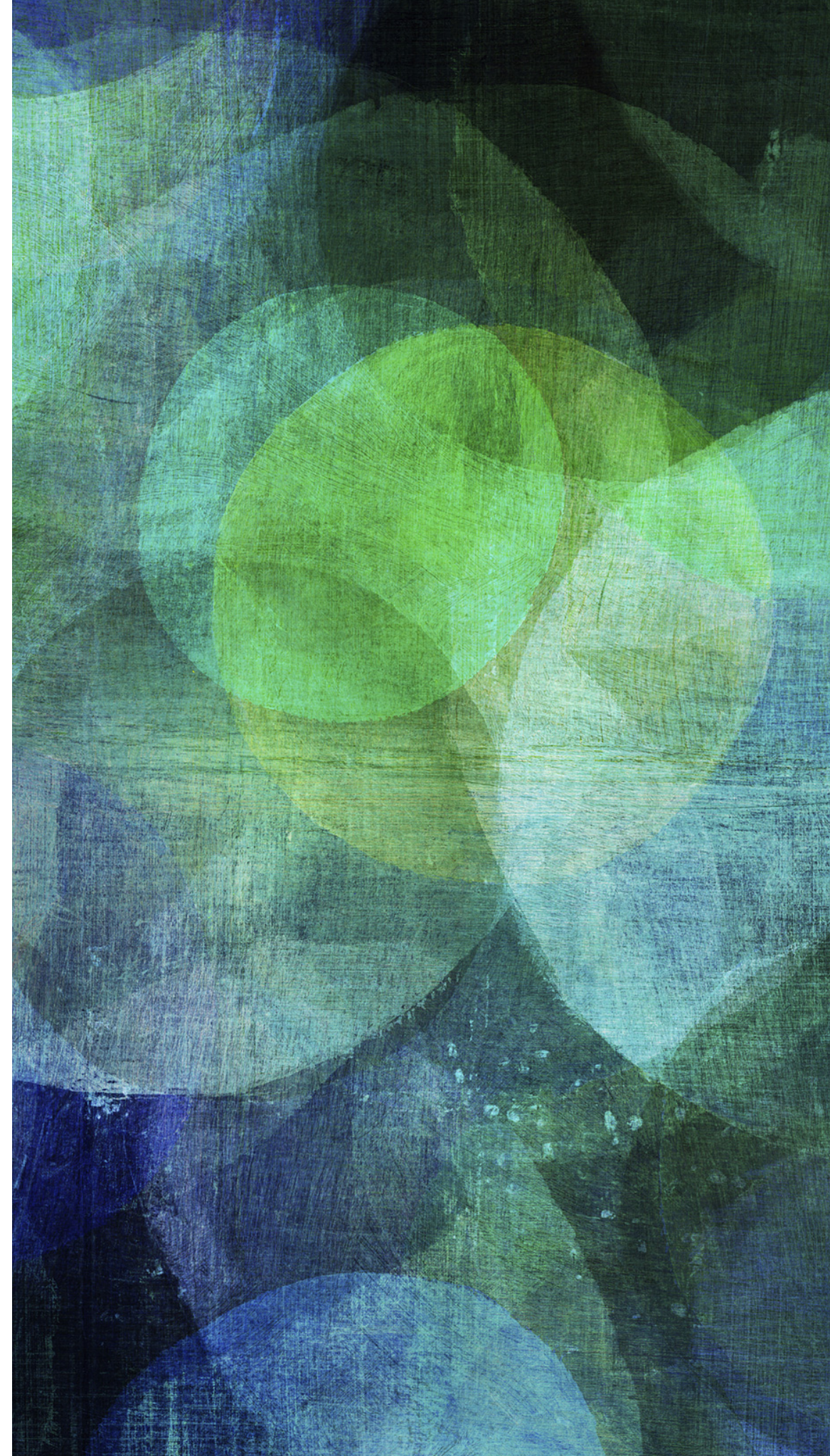


MÉTRICAS DE AVALIAÇÃO

Fundamentos de IA
Prof Carine G Webber

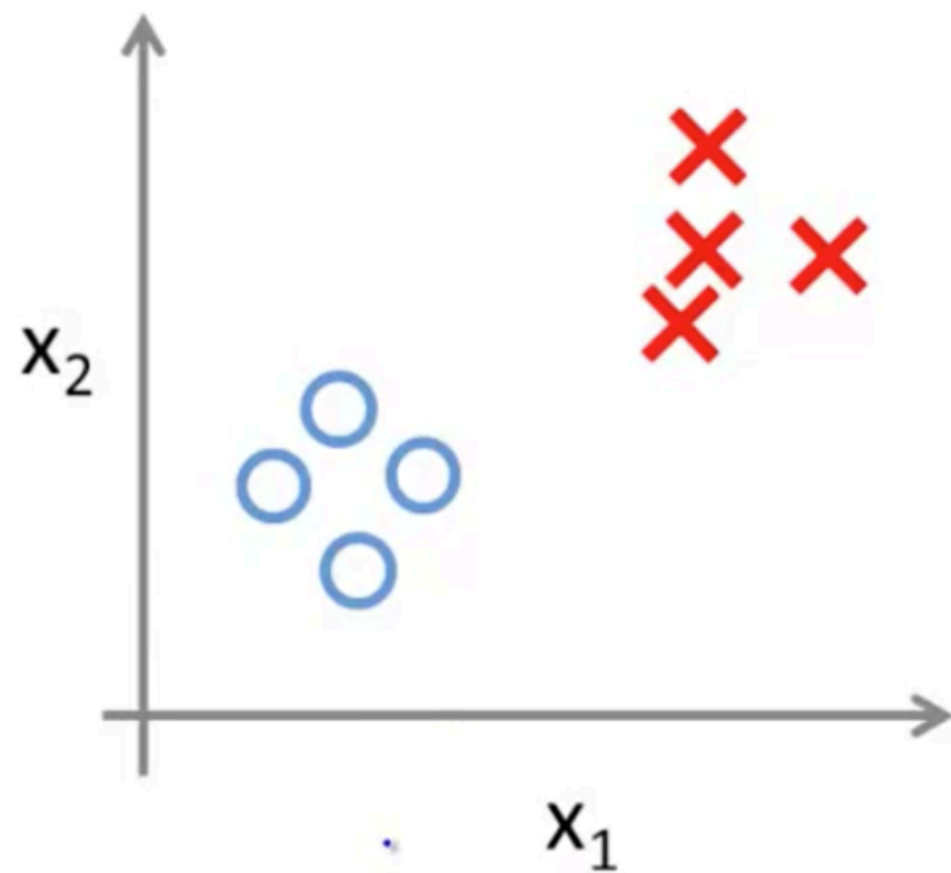


TERMINOLOGIA

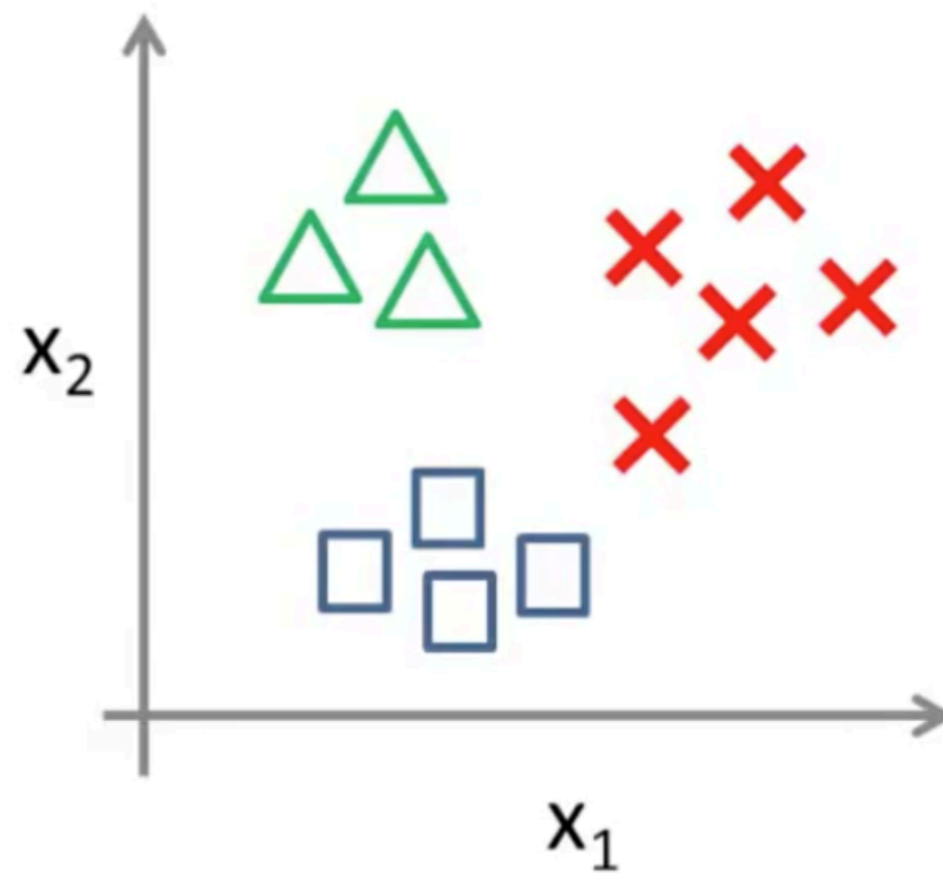
- Dataset: conjunto de dados, amostra
- Dados de treinamento: parte da amostra usada para construção do modelo
- Dados de teste: parte da amostra usada para teste do modelo
- Instância: um exemplar, caso ou indivíduo da amostra
- Classe: atributo alvo ou classificador
- Classificador multi-classe: o dataset contém 3 ou mais classes distintas.
- Classificador binário: o dataset contém 2 classes.

EXEMPLOS DE PROBLEMAS BINÁRIO X MULTI-CLASSE

Binário



Multi-classe



EXEMPLO DE DATASET BINÁRIO

Dataset Titanic : <https://www.kaggle.com/c/titanic/data>

Variable	Definition	Key
survival	Survival	0 = No, 1 = Yes
pclass	Ticket class	1 = 1st, 2 = 2nd, 3 = 3rd
sex	Sex	
Age	Age in years	
sibsp	# of siblings / spouses aboard the Titanic	
parch	# of parents / children aboard the Titanic	
ticket	Ticket number	
fare	Passenger fare	
cabin	Cabin number	
embarked	Port of Embarkation	C = Cherbourg, Q = Queenstown, S = Southampton

EXEMPLO DE DATASET MULTICLASSE

.....

<https://github.com/zalandoresearch/fashion-mnist>

<http://fashion-mnist.s3-website.eu-central-1.amazonaws.com/>



MATRIZ DE CONFUSÃO

Classificador Binário

		Detectada	
		Sim	Não
Real	Sim	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (VN)

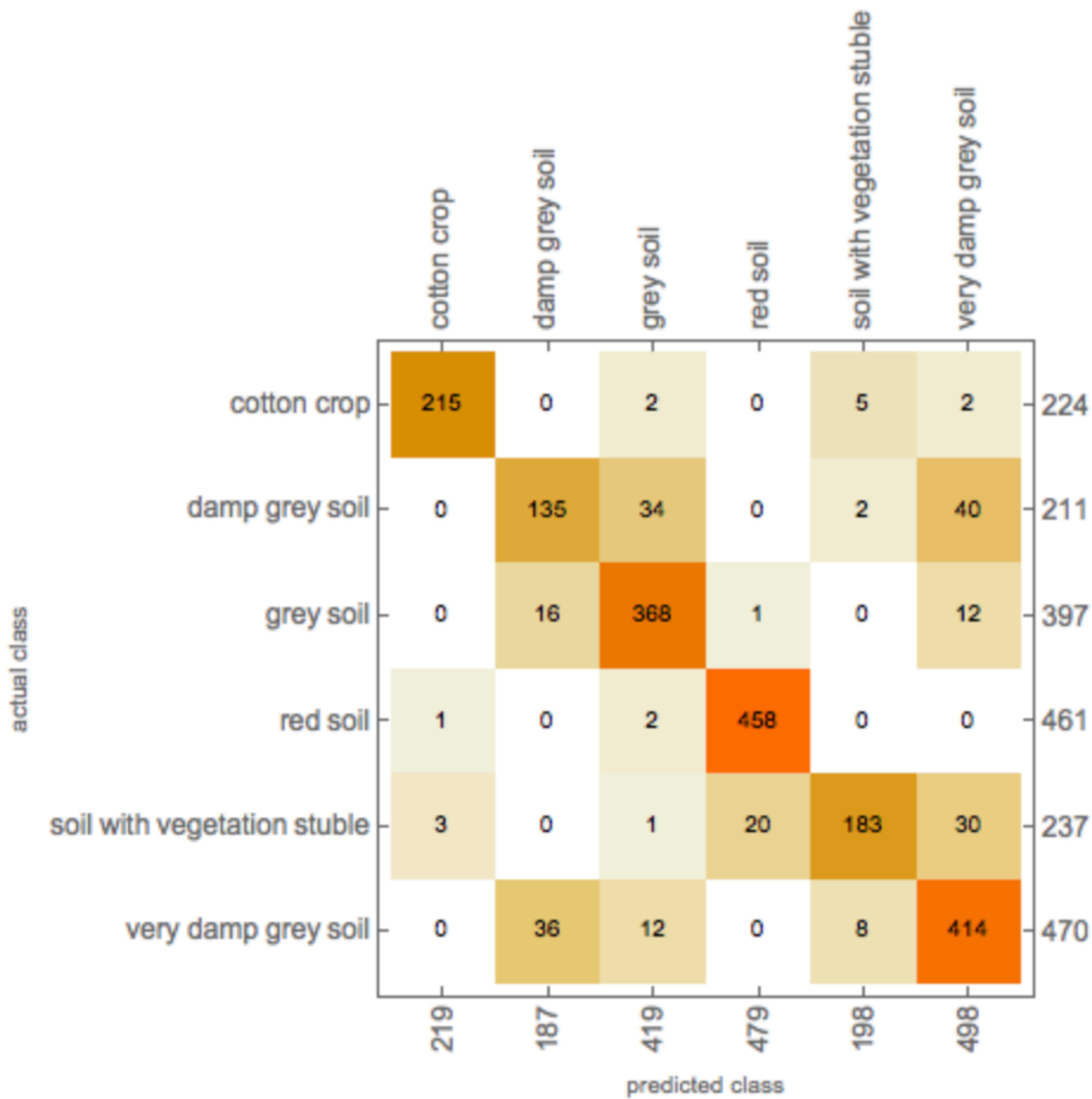
MATRIZ DE CONFUSÃO – CLASSIFICADOR BINÁRIO

.....

		Detectada	
		Sim	Não
Real	Sim	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (VN)

- Verdadeiros Positivos: classificação correta da classe Positivo;
- Falsos Negativos: erro em que o modelo previu a classe Negativo quando o valor real era classe Positivo;
- Falsos Positivos: erro em que o modelo previu a classe Positivo quando o valor real era classe Negativo;
- Verdadeiros Negativos: classificação correta da classe Negativo.

MATRIZ DE CONFUSÃO – CLASSIFICADOR MULTI-CLASSE



A diagonal principal apresenta a contagem de VP para cada Classe.

MÉTRICAS: ACURÁCIA

- A acurácia corresponde a porcentagem das instâncias dos dados de teste que são corretamente classificadas.
- $\text{Acurácia} = (\text{VP} + \text{VN}) / (\text{VP} + \text{FP} + \text{VN} + \text{FN})$
- $\text{Erro} = 1 - \text{Acurácia}$

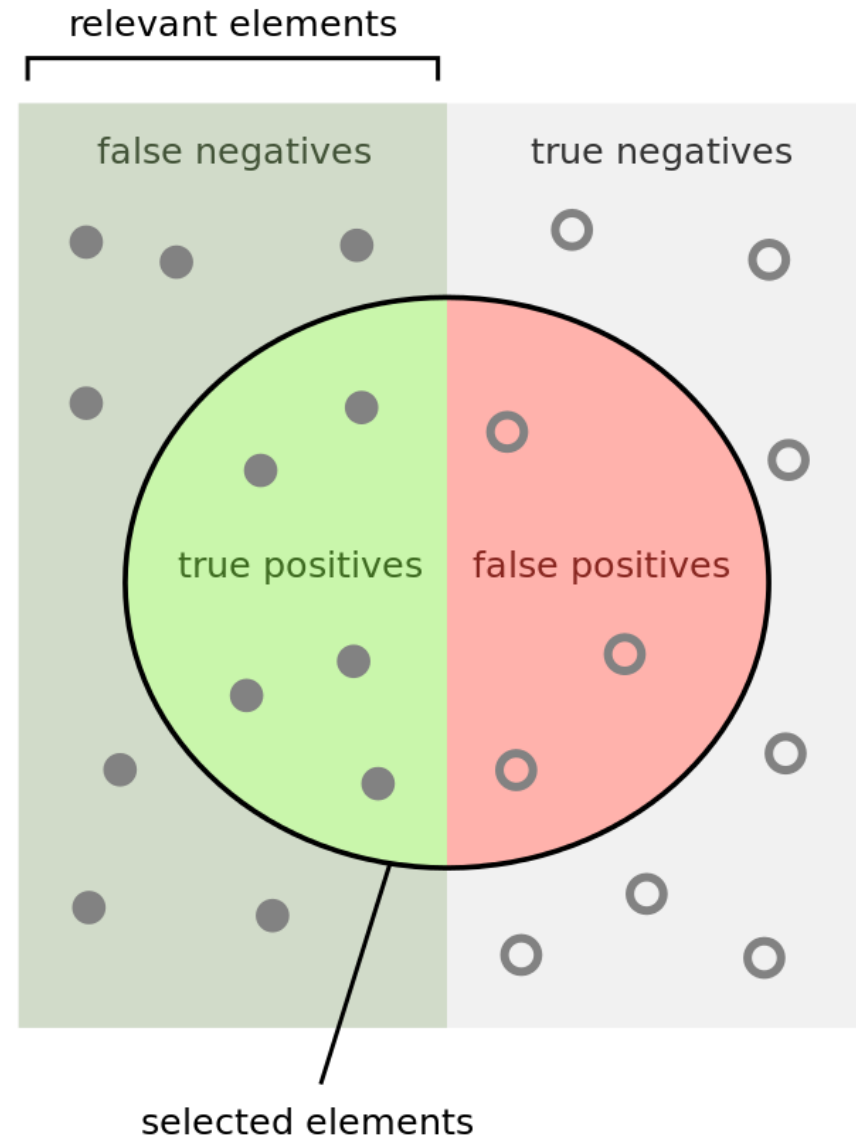
MÉTRICA: RECALL

- Sensibilidade : a proporção de verdadeiros positivos.
- Capacidade do sistema em prever corretamente a condição para casos que realmente a têm.
- $\text{Recall} = \text{VP} / (\text{VP} + \text{FN})$

MÉTRICA: PRECISÃO

- Calculada a proporção de verdadeiros positivos em relação a todas as predições positivas.
- Esta medida é altamente sensível a desbalanceamentos do conjunto de dados.
- $\text{Precisão} = \text{VP} / (\text{VP} + \text{FP})$

REPRESENTAÇÃO VISUAL



How many selected items are relevant?

$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

How many relevant items are selected?

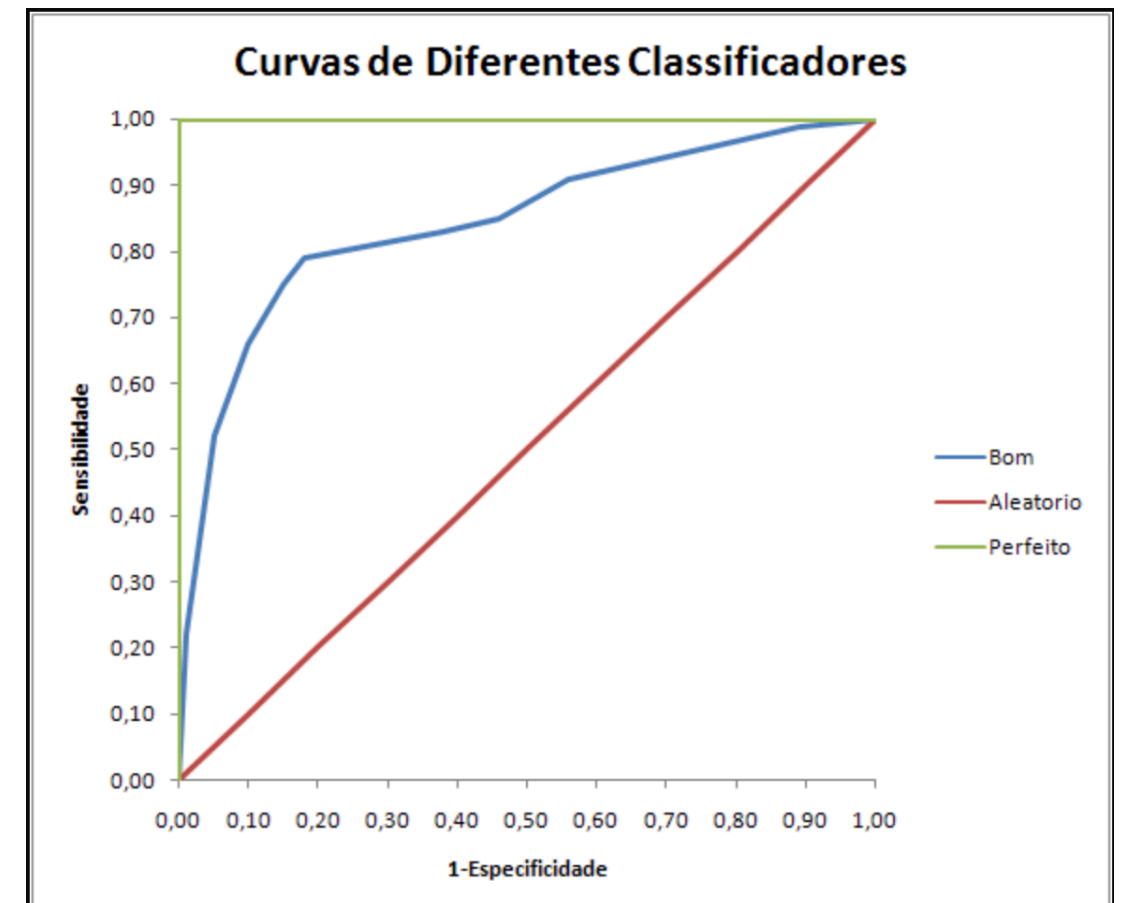
$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

F-MEASURE OU F-SCORE

- Métrica que pondera o recall e a precisão.
- Medida para avaliar globalmente o modelo.
- $F_measures = 2 * (recall * precisão) / (recall + precisão)$

CURVA ROC

- Receiver Operating Characteristics
- AUC - Area Under Curve
- Um classificador perfeito corresponde a uma linha horizontal no topo do gráfico, porém ela dificilmente será alcançada.
- Na prática, curvas consideradas boas estão entre a linha diagonal e a linha perfeita, pois quanto maior a distância da linha diagonal, melhor o classificador.



REFERÊNCIAS

- Jason Brownlee - <https://machinelearningmastery.com/about/>
- SABATTINI, R. M. E.; “Um Programa para o Cálculo da Acurácia, Especificidade e Sensibilidade de Testes Médicos”; Revista Informédica, 2 (12): 19-21, 1995.
Disponível em: <<http://www.informaticamedica.org.br/informed/sensib.htm>> Acesso em: 07 jul. 2009.
- ANAESTHETIST.COM, “Receiver Operating Curves: An Introduction”, Disponível em: <<http://www.anaesthetist.com/mnm/stats/roc/Findex.htm>> Acesso em: 13 jul. 2009.