

# Inteligência Artificial

## Aula 29 - Aprendizagem de Máquina: Medidas de Avaliação <sup>1</sup>

Sílvia M.W. Moraes

Faculdade de Informática - PUCRS

November 14, 2017

---

<sup>1</sup>Este material não pode ser reproduzido ou utilizado de forma parcial sem a permissão dos autores.

# Sinopse

- Nesta aula, apresentamos uma **visão geral de algumas medidas para avaliação de algoritmos de aprendizagem** - tarefas de **classificação e agrupamento**.

# Sumário

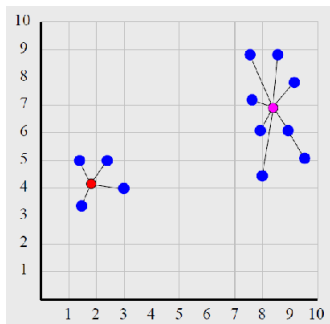
- 1 Medidas de Avaliação para Agrupamento
- 2 Medidas de Avaliação para Classificação

# Medida SSE

- Medida mais comum para a tarefa de agrupamento é a **Soma dos Erros Quadráticos** (SSE – sum of squared errors) : mede a coesão dos clusters
  - Para cada objeto, o erro é a distância ao grupo mais próximo (distância do objeto ao centróide do seu cluster).
  - Para obter SSE, elevamos os erros ao quadrado e os somamos.
  - $$SSE = \sum_{i=1}^k \sum_{x \in C_i} distancia^2(m_i, x)$$
  - onde  $k$  é o número de clusters,  $x$  é um objeto do grupo  $C_i$  e  $m_i$  é o centro do grupo  $C_i$ .

# Medida SSE

- Quanto menor o SSE, mais compactos (coesos) são os grupos, pois minimizar o SSE significa minimizar a variância intra-grupo.



# Medidas geradas pelo Weka

- Medidas geradas pelo Weka
  - **Correctly Classified Instances:** total de amostras corretamente classificadas e o seu percentual
  - **Incorrectly Classified Instances:** total de amostras incorretamente classificadas e o seu percentual
  - **Kappa statistic:** calcula o índice de concordância no processo de classificação (entre os avaliadores). Valores próximos a 1 indicam concordância, enquanto valores próximos a 0 indicam ausência de concordância além do puro acaso.

	Poor	Slight	Fair	Moderate	Substantial	Almost perfect
Kappa	0.0	.20	.40	.60	.80	1.0

# Medidas geradas pelo Weka

- Medidas geradas pelo Weka
  - **Mean absolute error** (erro absoluto médio): média do valor absoluto dos erros. Erro absoluto corresponde à diferença em módulo entre o valor esperado e sua aproximação.
  - **Root mean squared error** (raiz do erro quadrático médio): raiz da média dos quadrados dos erros. Erro quadrático corresponde ao quadrado da diferença entre o valor esperado e sua aproximação.
  - **Relative absolute error** (erro absoluto relativo) : corresponde à razão entre o erro absoluto e a magnitude do valor esperado.
  - **Root relative squared error** (erro quadrático relativo): corresponde à razão entre o erro quadrático e a magnitude do valor esperado.

# Medidas Usuais para Classificação

- As medidas mais usuais em tarefas de Classificação são
  - Accuracy (Acurácia),
  - Precision (Precisão)
  - Recall (Abrangência)
  - F-Measure (F1)
- São medidas clássicas definidas para a área de Recuperação de Informações.



# Classificação: Matriz de Confusão

- As medidas usadas para avaliar os resultados da classificação são calculadas a partir de uma **tabela de confusão** (ou de contingência).
- Nessa tabela são confrontados os resultados do preditor com os resultados esperados.
- É sempre uma matriz quadrada:  $nroClasses \times nroClasses$ , onde as linhas representam os resultados esperados e as colunas, os resultados gerados pelo preditor.

# Classificação: Matriz de Confusão

- As medidas foram originalmente definidas para classificadores binários, ou seja, capazes de classificar dados em apenas 2 classes. Para duas classes *Sim* e *Não*, a matriz seria assim:

		Preditor (Máquina)	
		Sim	Não
Resultados Esperados	Sim	TP	FP
	Não	FN	TN

- TP: Verdadeiro-Positivo** (classificação correta): era da classe esperada e o preditor o classificou nessa classe.
- TN: Verdadeiro-Negativo** (classificação correta): não era da classe e preditor não o classificou nessa classe.
- FP: Falso-positivo** (classificação incorreta): não era da classe esperada e o preditor classificou a amostra nessa classe.
- FN: Falso-negativo** (classificação incorreta): era da classe e o preditor classificou a amostra em outra classe.

# Classificação: Matriz de Confusão

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.

=== Confusion Matrix ===

```

  a  b  c  <-- classified as
40  0  0 |  a = Iris-setosa
 0 37  3 |  b = Iris-versicolor
 0  2 38 |  c = Iris-virginica

```

- A tabela corresponde a uma matriz quadrada  $3 \times 3$ , onde as linhas correspondem às classes esperadas (conhecidas, que estão no arquivo de teste) e as colunas às classes previstas.
- A diagonal principal corresponde aos acertos do classificador.
- **Para classe Iris-Versicolor:**
  - TP (Verdadeiro-Positivo): 37
  - TN (Verdadeiro-Negativo): 78
  - FP (Falso-Positivo): 2
  - FN (Falso-Negativo): 3

# Classificação: Matriz de Confusão

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.

=== Confusion Matrix ===

a	b	c	<-- classified as
40	0	0	a = Iris-setosa
0	37	3	b = Iris-versicolor
0	2	38	c = Iris-virginica

- **Atividade I:** Calcule os índices TP, TN, FP e FN para as classe Iris-Virginica e Iris-Setosa

# Classificação: Matriz de Confusão

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.

=== Confusion Matrix ===

```
a  b  c  <-- classified as
40  0  0 | a = Iris-setosa
  0 37  3 | b = Iris-versicolor
  0  2 38 | c = Iris-virginica
```

- **Atividade I:** Calcule os índices TP, TN, FP e FN para as classe Iris-Virginica e Iris-Setosa
- **Resposta:** Para classe **Iris-Virginica**:
  - TP (Verdadeiro-Positivo): 38
  - TN (Verdadeiro-Negativo): 77
  - FP (Falso-Positivo): 3
  - FN (Falso-Negativo): 2

# Classificação: Matriz de Confusão

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.

=== Confusion Matrix ===

```
a  b  c  <-- classified as
40  0  0 |  a = Iris-setosa
  0 37  3 |  b = Iris-versicolor
  0  2 38 |  c = Iris-virginica
```

- **Atividade I:** Calcule os índices TP, TN, FP e FN para as classe Iris-Virginica e Iris-Setosa
- **Resposta:** Para classe **Iris-Setosa**:
  - TP (Verdadeiro-Positivo): 40
  - TN (Verdadeiro-Negativo): 75
  - FP (Falso-Positivo): 0
  - FN (Falso-Negativo): 0

# Classificação: Accuracy

- Acurácia e precisão, embora, muitas vezes, sejam usadas como sinônimos, estatisticamente não são a mesma medida.
- **Accuracy**: corresponde à proporção de acertos do classificador.

		Preditor (Máquina)	
		Sim	Não
Resultados Esperados	Sim	TP	FP
	Não	FN	TN

- TP: Verdadeiro-Positivo (classificação correta)
- TN: Verdadeiro-Negativo (classificação correta)
- FP: Falso-positivo (classificação incorreta)
- FN: Falso-negativo (classificação incorreta)
- $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$

# Classificação: Accuracy

- Micro-médias (por classe)

- $Accuracy_{classe} = \frac{TP+TN}{TP+TN+FP+FN}$

- Macro-médias (índices gerais)

- *Accuracy Média*: média aritmética de todas as micro-médias de *accuracy*.



## Classificação: Accuracy

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.

- **Para classe Iris-Versicolor:**

- TP (Verdadeiro-Positivo): 37
- TN (Verdadeiro-Negativo): 78
- FP (Falso-Positivo): 2
- FN (Falso-Negativo): 3

- $Accuracy_{versicolor} = \frac{TP+TN}{TP+TN+FP+FN} = \frac{37+78}{37+78+2+3} = 0,958$

## Classificação: Accuracy

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.

=== Confusion Matrix ===

```

a  b  c  <-- classified as
40  0  0 | a = Iris-setosa
 0 37  3 | b = Iris-versicolor
 0  2 38 | c = Iris-virginica
    
```

- **Atividade II:** Calcule a accuracy (micro-média) para as classes iris-setosa e iris-virginica. Em seguida, calcula a macro-média de accuracy.

## Classificação: Accuracy

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.
- **Atividade II:** Calcule a accuracy (micro-média) para as classes iris-setosa e iris-virginica. Em seguida, calcula a macro-média de accuracy.
  - **Resposta:** Para classe Iris-Virginica:
    - TP (Verdadeiro-Positivo): 38
    - TN (Verdadeiro-Negativo): 77
    - FP (Falso-Positivo): 3
    - FN (Falso-Negativo): 2
  - $Accuracy_{virginica} = \frac{TP+TN}{TP+TN+FP+FN} = \frac{38+77}{38+77+3+2} = 0,958$

## Classificação: Accuracy

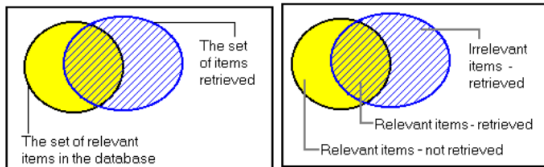
- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.
- **Atividade II:** Calcule a accuracy (micro-média) para as classes iris-setosa e iris-virginica. Em seguida, calcula a macro-média de accuracy.
  - **Resposta:** Para classe Iris-Setosa:
    - TP (Verdadeiro-Positivo): 40
    - TN (Verdadeiro-Negativo): 75
    - FP (Falso-Positivo): 0
    - FN (Falso-Negativo): 0
  - $Accuracy_{setosa} = \frac{TP+TN}{TP+TN+FP+FN} = \frac{40+75}{40+75} = 1$

## Classificação: Accuracy

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.
- **Atividade II:** Calcule a accuracy (micro-média) para as classes iris-setosa e iris-virginica. Em seguida, calcula a macro-média de accuracy.
  - **Resposta: Macro-Média:**  $Accuracy = \frac{1+0,958+0,958}{3} = 0,972$
  - $Accuracy_{setosa} = 1$
  - $Accuracy_{versicolor} = 0,958$
  - $Accuracy_{virginica} = 0,958$

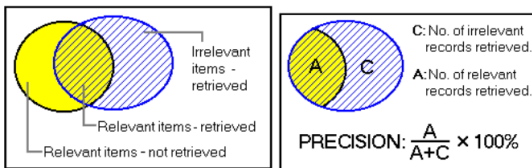
# Classificação: Precision e Recall

- Para entender a diferença entre Precision e Recall, considere as imagens abaixo.



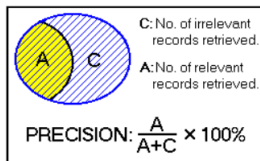
# Classificação: Precision

- **Precision:** é a razão entre número de dados corretamente classificados e o número total dados classificados (corretamente e incorretamente).



# Classificação: Precision

- **Precision:** proporção de textos corretamente classificados.



		Preditor (Máquina)	
		Sim	Não
Resultados Esperados	Sim	TP	FP
	Não	FN	TN

*Precision =*

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



# Classificação: Precision

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.
  - Para classe **Iris-Versicolor**:
    - TP (Verdadeiro-Positivo): 37
    - TN (Verdadeiro-Negativo): 78
    - FP (Falso-Positivo): 2
    - FN (Falso-Negativo): 3

$$Precision_{versicolor} = \frac{TP}{TP+FP} = \frac{37}{37+2} = 0,9487$$

# Classificação: Precision

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.
  - **Atividade III:** Calcule a precision (micro-média) para as classes iris-setosa e iris-virginica. Em seguida, calcula a macro-média de precision.

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

# Classificação: Precision

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.
  - **Atividade III:** Calcule a precision (micro-média) para as classes iris-setosa e iris-virginica. Em seguida, calcula a macro-média de precision.
  - **Resposta:** Para classe **Iris-Setosa**:
    - TP (Verdadeiro-Positivo): 40
    - TN (Verdadeiro-Negativo): 75
    - FP (Falso-Positivo): 0
    - FN (Falso-Negativo): 0

$$Precision_{setosa} = \frac{TP}{TP+FP} = \frac{40}{40+0} = 1$$

# Classificação: Precision

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.
  - **Atividade III:** Calcule a precision (micro-média) para as classes iris-setosa e iris-virginica. Em seguida, calcula a macro-média de precision.
  - **Resposta:** Para classe **Iris-Virginica**:
    - TP (Verdadeiro-Positivo): 38
    - TN (Verdadeiro-Negativo): 77
    - FP (Falso-Positivo): 3
    - FN (Falso-Negativo): 2

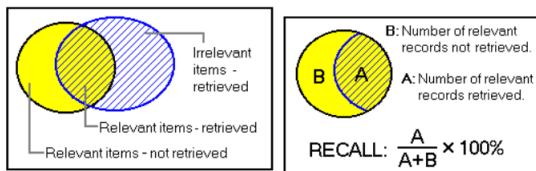
$$Precision_{virginica} = \frac{TP}{TP+FP} = \frac{38}{38+3} = 0,9268$$

# Classificação: Precision

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.
  - **Atividade III:** Calcule a precision (micro-média) para as classes iris-setosa e iris-virginica. Em seguida, calcula a macro-média de precision.
  - **Resposta: Macro-Média:**  $Precision = \frac{1+0,9268+0,9487}{3} = 0,9585$ 
    - $Precision_{setosa} = \frac{TP}{TP+FP} = \frac{40}{40+0} = 1$
    - $Precision_{virginica} = \frac{TP}{TP+FP} = \frac{38}{38+3} = 0,9268$
    - $Precision_{versicolor} = \frac{TP}{TP+FP} = \frac{37}{37+2} = 0,9487$

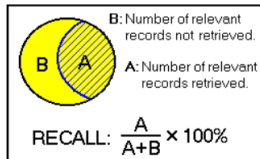
# Classificação: Recall

- **Recall:** corresponde à razão entre de dados classificados corretamente (relevantes) sobre o número total de classificações definidas como corretas.



# Classificação: Recall

- Recall:** proporção de textos que deveriam ter sido corretamente classificados.



		Preditor (Máquina)	
		Sim	Não
Resultados Esperados	Sim	TP	FP
	Não	FN	TN

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

# Classificação: Recall

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.
  - Para classe **Iris-Versicolor**:
    - TP (Verdadeiro-Positivo): 37
    - TN (Verdadeiro-Negativo): 78
    - FP (Falso-Positivo): 2
    - FN (Falso-Negativo): 3

$$Recall_{Versicolor} = \frac{TP}{TP+FN} = \frac{37}{37+3} = 0,925$$



## Classificação: Recall

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.
  - **Atividade IV:** Calcule o recall (micro-média) para as classes iris-setosa e iris-virginica. Em seguida, calcula a macro-média de recall.

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

# Classificação: Recall

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.
  - **Atividade IV:** Calcule o recall (micro-média) para as classes iris-setosa e iris-virginica. Em seguida, calcula a macro-média de recall.
  - **Resposta:** Para classe **Iris-Virginica**:
    - TP (Verdadeiro-Positivo): 38
    - TN (Verdadeiro-Negativo): 77
    - FP (Falso-Positivo): 3
    - FN (Falso-Negativo): 2

$$Recall_{virginica} = \frac{TP}{TP+FN} = \frac{38}{38+2} = 0,95$$

# Classificação: Recall

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.
  - **Atividade IV:** Calcule o recall (micro-média) para as classes iris-setosa e iris-virginica. Em seguida, calcula a macro-média de recall.
  - **Resposta:** Para classe Iris-Setosa:
    - TP (Verdadeiro-Positivo): 40
    - TN (Verdadeiro-Negativo): 75
    - FP (Falso-Positivo): 0
    - FN (Falso-Negativo): 0

$$Recall_{setosa} = \frac{TP}{TP+FN} = \frac{40}{40+0} = 1$$

## Classificação: Recall

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.
  - **Atividade IV:** Calcule o recall (micro-média) para as classes iris-setosa e iris-virginica. Em seguida, calcula a macro-média de recall.
  - **Resposta: Macro-Média:**  $Recall = \frac{1+0,95+0,925}{3} = 0,9583$ 
    - $Recall_{setosa} = \frac{TP}{TP+FN} = \frac{40}{40+0} = 1$
    - $Recall_{virginica} = \frac{TP}{TP+FN} = \frac{38}{38+2} = 0,95$
    - $Recall_{Versicolor} = \frac{TP}{TP+FN} = \frac{37}{37+3} = 0,925$

# Classificação: F-Measure

- Média Harmônica das medidas Precision e Recall

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

- **Exemplo:** 120 amostras das classes **Iris-Setosa(a)**, **Iris-Versicolor(b)** e **Iris-Virginica(c)** foram testadas, 40 de cada classe.

- $F - \text{Measure}_{\text{versicolor}} = \frac{2 \times 0,9487 \times 0,925}{0,9487 + 0,925} = 0,9367$
- $F - \text{Measure}_{\text{setosa}} = \frac{2 \times 1 \times 1}{1 + 1} = 1$
- $F - \text{Measure}_{\text{virginica}} = \frac{2 \times 0,9268 \times 0,95}{0,9268 + 0,95} = 0,9382$
- Macro-Média:  $F - \text{Measure} = \frac{0,9367 + 1 + 0,9382}{3} = 0,9583$