**Ciência de dados Regressão linear**

**Renato Moraes Silva renato.silva@facens.br**

**Cálculo diferencial Derivada de uma**

****

**Soma / Substração** *dx*(*u ± v*) = *du*

**constante** *d*

*dx*(*c*) = 0

*d*

*dx±dv dx*

**Derivada da potência** *d*

*dx*(*xn*) = *nxn−*1.

Portanto:

*d*

*dx*(*x*) = 1

**Produto por uma constante**

*dx*(*cv*) = *cdv*

*d*

*dx*

**Derivada do produto** *d*

*dx*(*uv*) = *udv*

*dx*+ *vdu*

*dx*

*Cálculo diferencial: exemplos* 

**Cálculo diferencial: exemplos**

Função constante Função linear Função quadrática

�� �� = 50

��′ �� =���� ����= 0

�� �� = 3�� + 1 ��′ �� =����

����= 3

�� �� = ��2

��′ �� =����

����= 2��

Facens - Especialização em Inteligência Artificial Aplicada 9

**Cálculo diferencial: exercícios **1 *f*(*x*) = 11

2 *f*(*x*) = 4*x* + 20*x* + 2

3 *f*(*x*) = 4*x*3 + *x* + 1

4 Se a população mundial é dada pela função *f*(*x*), qual interpretação pode ser dada à sua derivada?

**Derivadas parciais** 

Em funções de múltiplas variáveis, é possível estabelecer a de múltiplas variáveis, é possível estabelecer a derivada em relderivada em relação a cada variável particular

�� =��2ℎ��

3

Derivada em

relação à “r” Derivada em ����

����=2��ℎ�� 3

��ℎ=��2��3 ����

relação à “h”

**Ciência de dados Revisão de Álgebra Linear**

**Renato Moraes Silva**

**renato.silva@facens.br**

**Álgebra linear **Multiplicação por uma constante

3 *·* 2 4 7= 3 *·* 2 3 *·* 4 3 *·* 7= 6 12 21

1

2*·*



4 6 10 1



 =



4 *·*126 *·*12 10 *·*121 *·*12



 =



2 3 512

 

**Álgebra linear – multiplicação de matri** *lgebra linear* 

**zes**

Multiplicação exemplo:

*A × B*: a quantidade de colunas de A deve ser igual a quantidade

Facens - Especialização em Inteligência Artificial Aplicada 14

de linhas de B

O produto de uma matriz A com dimensão *m × n* e uma matriz B com dimensão *n × k*, resulta em uma matriz *m × k*

**Álgebra linear: exemplos em Numpy** *Álgebra linear: Exemplos em Numpy* 

Facens - Especialização em Inteligência Artificial Aplicada 15

**Exercícios **

1 Suponha que você irá multiplicar uma matriz de dimensão 3x2 por uma de 2x3. É possível fazer essa multiplicação? Qual será a dimensionalidade da nova matriz?

2 Suponha que você irá multiplicar uma matriz de dimensão 4x2 por uma de 4x3. É possível fazer essa multiplicação? Qual será a dimensionalidade da nova matriz?

3 Considere *A* =

1 2 3 2

e *B* =

1 2 3 6 4 5

. Quanto é *A × B*?

4 Considere *A* =

1 2 3 1 2 3

e *B* =

 

1 2



. Quanto é *A × B*?

5 Calcule: 5 *×*

1 2 2 1

3

**Ciência de dados Regressão**

**Renato Moraes Silva renato.silva@facens.br**

**Modelos preditivos**Tiago A. Almeida Modelos preditivos 

Definição: dado conjunto de observações

**D** = {(**x**(i), *f*(**x**(i))), *i* = 1, ..., *m*}

*f* representa uma função desconhecida: função objetivo

Mapeia entradas em saídas correspondentes

**^**

Algoritmo preditivo aprende aproximação *f*

Permite estimar valor de *f* para novos objetos **x**

Regressão

*y* (i) = *f*(**x**(i)) ∈ ℜ **^**

**Regressão**

Tiago A. Almeida

Regressão 

Meta: aprender função (curva aproximada) que relacione entradas a valores contínuos de saídas

**Exemplos**

Prever valor de mercado de um imóvel

Prever o lucro de um empréstimo bancário

Prever o valor de um ativo

**Outros exemplos **

Qual é o valor comercial de uma casa, dadas as suas características e de seu bairro?

Qual é o salário de um determinado profissional, dado o número de anos em que ele está na profissão?

Qual é o valor adequado para emprestar de dinheiro a um determinado cliente de um banco?

Quanto uma determinada região da cidade produz de lixo de acordo com sua densidade populacional?

**Pipeline**Pipeline 

conjunto de dados

*x*11 *x*12 ... Técnica de

*x*21 *x*22 ...

*x*1n *x*2n

*y*1 *y*2

*h*

...... .. ...

regressão

Hipótese

*x*m1 *x*m2 ...

*x*mn

*y*m

**Treinamento**

*Objetivo:* preditor capaz de predizer corretamente o valor de novos dados

xn x ... **x** x1 2 novo dado

*h*

Preditor

*h*(**x**)

previsão **Teste**

**Métodos**

Regressão Linear

Tiago A. Almeida

Métodos 

Árvores de Regressão

Redes Neurais Artificiais

Máquinas de Vetores de Suporte

… Cada técnica de AM utiliza

um mecanismo distinto na

aproximação da função

objetivo

**Ciência de dados**

**Regressão Linear Univariada**

**Renato Moraes Silva**

**renato.silva@facens.br**

**Regressão: exemplo**



o de treinamento (valores de imóveis) Conjunto de treinamento (valores de imóveis)

Valor

(em US$ 1000)

500"

400"

300"

200"

100"

0"

0" 500" 1000" 1500" 2000" 2500" 3000" Tamanho

egresso: exempo **Regressão: exemplo** 

Conjunto de treinamento (valores de imóveis) Conjunto de treinamento (valores de imóveis)

**Notação**

**m** = número de amostras **x** = atributos (entrada) **y** = meta (saída)

**Tamanho (x) Valor x U$ 1000 (y)** 2104 460

1416 232

1534 315

852 178

... ...

**Hipótese (h)**Tiago A. Almeida Hipótese (h) 

y

x

**Hipótese (h)**Tiago A. Almeida Hipótese (h) 

Hipótese:

*h✓*(*x*) = *✓*0 + *✓*1(*x*)

*h*(*x*)

y

x

- Regressão linear com uma variável - Regressão linear univariada

**Hipótese (h) **

*y* ou *hθ*(*x*) é também chamada de variável dependente *x* é chamada de variável independente

*θ*0 determina onde a reta cruza o eixo y

*θ*1 define a inclinação da reta

**Hipótese (h)** Hipótese:

Tiago A. Almeida

Hipótese (h) *h✓*(*x*) = *✓*0 + *✓*1(*x*)

*h*(*x*)

y

x

**Como encontrar a melhor hipótese?**

- Regressão linear com uma variável - Regressão linear univariada

**Escolha da hipótese**Tiago A. Almeida Escolha da hipótese 

Conjunto de treinamento (valores de imóveis)

**Tamanho (x) Valor x U$ 1000 (y)**

2104 460

1416 232

1534 315

852 178

... ...

Hipótese:

*h✓*(*x*) = *✓*0 + *✓*1(*x*)

Parâmetros:

*✓i*

*✓i*

Como escolher bons valores para ?

**Parâmetros**

Tiago A. Almeida

Parâmetros 

Hipótese: *h✓*(*x*) = *✓*0 + *✓*1(*x*)

3"

2"

1"

0"

0" 1" 2" 3"

3"

2"

1"

0"

0" 1" 2" 3"

3"

2"

1"

0"

0" 1" 2" 3"

**Parâmetros**

Tiago A. Almeida

Parâmetros 

Hipótese: *h✓*(*x*) = *✓*0 + *✓*1(*x*)

3" 2" 1" 0"

*h*(x) = 1.5 + 0x

0" 1" 2" 3"

3"

2"

1"

0"

0" 1" 2" 3"

3"

2"

1"

0"

0" 1" 2" 3"

**Parâmetros**

Tiago A. Almeida

Parâmetros 

Hipótese: *h✓*(*x*) = *✓*0 + *✓*1(*x*)

3"

3"

3"

*h*(x) = 1.5 + 0x *h*(x) = 0 + 0.5x 2"

2"

2"

1"

1"

1"

0"

0"

0"

0" 1" 2" 3"

0" 1" 2" 3"

0" 1" 2" 3"

**Parâmetros**

Tiago A. Almeida

Parâmetros 

Hipótese: *h✓*(*x*) = *✓*0 + *✓*1(*x*)

3"

3"

3"

*h*(x) = 1.5 + 0x *h*(x) = 0 + 0.5x *h*(x) = 1 + 0.5x

2"

1"

0"

0" 1" 2" 3"

2"

1"

0"

0" 1" 2" 3"

2"

1"

0"

0" 1" 2" 3"

**Parâmetros**

Tiago A. Almeida

Parâmetros 

Hipótese: *h✓*(*x*) = *✓*0 + *✓*1(*x*)

*h*(*x*)

y

x

*✓*0*, ✓*1

**Objetivo:** encontrar de tal forma que *h*(**x**) aproxima ***y*** para os exemplos de treinamento.

**Parâmetros**

Tiago A. Almeida

Parâmetros 

Hipótese: *h✓*(*x*) = *✓*0 + *✓*1(*x*)

y

x

*h*(*x*)

*✓*0*, ✓*1

**Em outras palavras:**

*✓*0*, ✓*1

- Escolher que minimize o **Erro Quadrático Médio** entre as predições (*h*(**x**)) e as saídas esperadas (***y***) para as amostras de treinamento.

**Objetivo:** encontrar de tal forma que *h*(**x**) aproxima ***y*** para os exemplos de treinamento.

Tiago A. Almeida

**Função custo (J)** 

Função Custo (J)

Erro quadrático médio: *h✓*(*x*) = *✓*0 + *✓*1(*x*)

y

x

*J*(*✓*0*, ✓*1) = 12*m*X*m i*=1

(*h✓*(*xi*) *yi*)2

**Função custo (J)**Tiago A. Almeida Função Custo (J) 

Minimizar o erro:

*h✓*(*x*) = *✓*0 + *✓*1(*x*)

*J*(*✓*0*, ✓*1) = 12*m*X*m i*=1

(*h✓*(*xi*) *yi*)2

y

x

*✓*0*, ✓*1

*min J*(*✓*0*, ✓*1) *✓*0*, ✓*1

**Objetivo:** encontrar de tal forma que *h*(**x**) aproxima ***y*** para os exemplos de treinamento.

**Função custo (J): interpretação**Tiago A. Almeida Função Custo (J): Interpretação 

Hipótese:

*h✓*(*x*) = *✓*0 + *✓*1(*x*)

Parâmetros:

*✓*0*, ✓*1

Função Custo:

Objetivo:

*J*(*✓*0*, ✓*1) = 12*m*X*m i*=1

*min J*(*✓*0*, ✓*1) *✓*0*, ✓*1

(*h✓*(*xi*) *yi*)2

**Função custo (J): interpretação**Tiago A. Almeida Função Custo (J): Interpretação

Hipótese: Parâmetros:

*h✓*(*x*) = *✓*0 + *✓*1(*x*) *✓*0*, ✓*1

Vamos supor que:

Função Custo:

*J*(*✓*0*, ✓*1) = 12*m*X*m*

*i*=1

*min J*(*✓*0*, ✓*1)

(*h✓*(*xi*) *yi*)2

Objetivo:

*min J*(*✓*1) *✓*0*, ✓*1 *✓*1

Tiago A. Almeida

**Função custo (J): interpretação** 

Função Custo (J): Interpretação

y

3" 2" 1" 0"

3" 2" 1" 0"

**X**

0" 1" 2" 3" x

&0,5" 0" 0,5" 1" 1,5" 2" 2,5"

**Função custo (J): interpretação**Tiago A. Almeida Função Custo (J): Interpretação

y

3" 2" 1"

3" 2" 1" 0"

**X**

0" **X**

0" 1" 2" 3" x

&0,5" 0" 0,5" 1" 1,5" 2" 2,5"

**Função custo (J): interpretação**Tiago A. Almeida Função Custo (J): Interpretação

y

3" 2" 1" 0"

3" 2" 1" 0"

**X**

**X**

**X**

0" 1" 2" 3" x

&0,5" 0" 0,5" 1" 1,5" 2" 2,5"

**Função custo (J): interpretação**Tiago A. Almeida Função Custo (J): Interpretação 

Variando obtém-se

uma parábola para

3"

Como deseja-se obter que minimize , então = 1 é a melhor escolha

2" 1" 0"

**X**

**X**

**X**

**X**

**X**

para esse exemplo, pois

&0,5" 0" 0,5" 1" 1,5" 2" 2,5"

Tiago A. Almeida

**Função custo (J): interpretação** 

Função Custo (J): Interpretação

Variando obtém-se

uma parábola para

3"

Como deseja-se obter que minimize , então = 1 é a melhor escolha

2" 1" 0"

**X**

**X**

**X**

**X**

**X**

para esse exemplo, pois

&0,5" 0" 0,5" 1" 1,5" 2" 2,5"

*J*(1) = 0

**Função custo (J): interpretação**Tiago A. Almeida Função Custo (J): Interpretação

Conjunto de treinamento (valores de imóveis)

Valor

(em US$

1000)

*✓*0 = 50

*✓*1 = 0*.*06

500"

400"

300"

200"

100"

0"

0" 500" 1000" 1500" 2000" 2500" 3000" Tamanho

**Função custo (J): interpretação**Tiago A. Almeida Função Custo (J): Interpretação

*J*(*✓*0*, ✓*1)

**Função custo (J): interpretação**Tiago A. Almeida Função Custo (J): Interpretação

500"

400"

300"

200"

100"

0"

0" 500" 1000" 1500" 2000" 2500" 3000"

Tiago A. Almeida

**Função custo (J): interpretação** 

Função Custo (J): Interpretação

500"

400"

300"

200"

100"

0"

0" 500" 1000" 1500" 2000" 2500" 3000"

**Função custo (J): interpretação**Tiago A. Almeida Função Custo (J): Interpretação

500"

400"

300"

200"

100"

0"

0" 500" 1000" 1500" 2000" 2500" 3000"

**Ciência de dados Gradiente Descendente**

**Renato Moraes Silva**

**renato.silva@facens.br**

Tiago A. Almeida

Gradiente descendente 

**Gradiente descendente**

)

1

θ

,

0

θ

(

J

θ0θ1

Tiago A. Almeida

Gradiente descendente 

**Gradiente descendente**

) 1

θ

,

0

θ

(J

**X**

θ0θ1

Tiago A. Almeida

Gradiente descendente 

**Gradiente descendente**

) 1

θ

,

0

θ

(J

**X**

**X**

θ0θ1

Tiago A. Almeida

Gradiente descendente 

**Gradiente descendente**

) 1

θ

,

0

θ

(J

**X**

**X**

**X**

θ0θ1

Tiago A. Almeida

Gradiente descendente 

**Gradiente descendente**

) 1

θ

,

0

θ

(J

**X**

**X**

**X**

**X**

θ0θ1

Tiago A. Almeida

Gradiente descendente 

**Gradiente descendente**

) 1

θ

,

0

θ

(J

**X**

**X**

**X**

**X**

**X**

θ0θ1

Tiago A. Almeida

Gradiente descendente 

**Gradiente descendente**

Repetir até convergir {

*✓j* := *✓j ↵@@✓jJ*(*✓*0*, ✓*1)

(para *j* = 0 e *j* = 1)

A atualização dos deve ser “simultânea” *✓j*

}

Tiago A. Almeida

Gradiente descendente 

**Gradiente descendente**

Repetir até convergir {

*✓j* := *✓j ↵@@✓jJ*(*✓*0*, ✓*1)

(para *j* = 0 e *j* = 1)

} Taxa de  aprendizagem

Derivada parcial (tamanho do passo)

Tiago A. Almeida

Gradiente descendente 

**Gradiente descendente**

Repetir até convergir {

*✓j* := *✓j ↵@@✓jJ*(*✓*0*, ✓*1)

(para *j* = 0 e *j* = 1)

- Se α for muito pequeno, a convergência poderá

ser lenta

- Se α for muito grande, poderá não haver

}

Taxa de

convergência

Derivada parcial

aprendizagem

(tamanho do passo)

Tiago A. Almeida

Gradiente descendente 

**Gradiente descendente**

Repetir até convergir {

*✓j* := *✓j ↵@@✓jJ*(*✓*0*, ✓*1)

(para *j* = 0 e *j* = 1)

- O método converge para um ótimo local mesmo

se α for um valor fixo

- Conforme o ótimo for se aproximando, o método

}

Taxa de

reduz o tamanho do passo automaticamenteDerivada parcial

aprendizagem

(tamanho do passo)

**Gradiente para regressão linear**Tiago A. Almeida Gradiente para regressão linear 

Repetir até convergir {

*✓j* := *✓j ↵@@✓jJ*(*✓*0*, ✓*1)

(para *j* = 0 e *j* = 1)

} *J*(*✓*0*, ✓*1) = 12*m*X*m i*=1

(*h✓*(*xi*) *yi*)2

*h✓*(*x*) = *✓*0 + *✓*1(*x*)

**Gradiente para regressão linear**Tiago A. Almeida Gradiente para regressão linear 

Repetir até convergir {

*✓j* := *✓j ↵@@✓jJ*(*✓*0*, ✓*1)

(para *j* = 0 e *j* = 1)

*j* =0: *@@✓jJ*(*✓*0*, ✓*1) =

}

*j* =1: *@@✓jJ*(*✓*0*, ✓*1) =

1

*m*

1

*m*

X*m*

*i*=1 X*m*

*i*=1

(*h✓*(*xi*) *yi*) (*h✓*(*xi*) *yi*)*xi*

**Derivadas da função de custo Erro Médio Quadrático (MSE)**

****

1 A Função de custo: *J* =1*m*X*m i*=1

(*yi − hθ*(*x*))2

2 A derivada em relação ao *θ*:*∂J*

*∂θ*1?

1 Expandindo o cálculo da predição:

*J* =1*m*X*m*

*i*=1

2 Distribuindo o expoente: h

[*yi −* (*xiθ*1 + *θ*0)]2

*J* =1*m*X*m i*=1

*y*2*i −*2*yi* (*xiθ*1 + *θ*0)+ (*−xiθ*1 *− θ*0)2i

**Derivadas da função de custo Erro Médio Quadrático (MSE)**

3 Distributiva no segundo termo e no expoente do terceiro:



*J* =1*m*X*m i*=1

h

*y*2*i −*2*yixiθ*1 *−* 2*yiθ*0+

(*−xiθ*1)2 + 2*xiθ*1*θ*0 + (*−θ*0)2 i

4 Calculando derivadas parciais para cada elemento. Lembrem que elementos onde não aparece *θ*1 são considerados constantes, logo suas derivadas são 0. Onde há *θ*1 a derivada torna-se 1, e onde há *θ*21, a derivada torna-se 2*θ*1:

*∂θ* =1*m*X*m*

*∂J*

*i*=1

*~~y~~*2*i −*2*yixi~~−~~*~~2~~*~~y~~~~i~~~~θ~~*~~0~~~~+~~2*x*2*iθ*1 + 2*xiθ*0*~~−θ~~*20

**Derivadas da função de custo Erro Médio Quadrático (MSE)**

****

5 Reorganizando os termos por fator comum (colocando *−*2*xi* em evidência):

*∂θ* =1*m*X*m*

*∂J*

*i*=1

[*−*2*xi* (*yi −* (*xiθ*1 + *θ*0))]

**Derivadas da função de custo Erro Médio Quadrático (MSE)**

3 A derivada em relação ao *θ*0:*∂J*

*∂θ*0?

1 Partindo da função J já com a distributiva aplicada:



*J* =1*m*X*m i*=1

h

*y*2*i −* 2*yixiθ*1 *−* 2*yiθ*0 +

(*−xiθ*1)2 + 2*xiθ*1*θ*0 + (*−θ*0)2 i

2 Calculando derivadas parciais para cada elemento. Lembrem que elementos onde não aparece *θ*0 são considerados constantes, logo suas derivadas são 0. Onde há *θ*0 a derivada torna-se 1, e onde há *θ*20. a derivada torna-se 2*θ*0:

*∂θ*0=1*m*X*m*

*∂J*

*i*=1

h

*~~y~~*2*i ~~−~~*~~2~~*~~y~~~~i~~~~x~~~~i~~~~θ~~*~~1~~ *−* 2*yi* +

*−*~~(2~~*~~x~~~~i~~~~θ~~*~~1~~~~)~~2 + 2*xiθ*1 + 2*θ*0 i

**Derivadas da função de custo Erro Médio Quadrático (MSE)**

****

3 Reorganizando os termos por fator comum (colocando *−*2 em evidência):

*∂θ* =1*m*X*m*

*∂J*

*i*=1

[*−*2 (*yi − xiθ*1 *− θ*0)]

*∂θ* =1*m*X*m*

*∂J*

*i*=1

[*−*2 (*yi −* (*xiθ*1 + *θ*0))]

**Gradiente para regressão linear**Tiago A. Almeida Gradiente para regressão linear**Algoritmo** 

repetir até convergir {

*✓*0 := *✓*0 *↵* 1*m*X*m*

*i*=1

*✓*1 := *✓*1 *↵* 1*m*X*m*

*i*=1

(*h✓*(*xi*) *yi*) (*h✓*(*xi*) *yi*)*xi*

Atualizar “simultaneamente” e θ0 θ1

}

Tiago A. Almeida

Gradiente descendente 

**Gradiente descendente**

**Observações**

GD é sensível à escala dos dados

É recomendável aplicar normalização por padronização para fazer com que

GD é sensível à taxa de aprendizagem

Recomendável testar com ..., 0.001, 0.01, 0.1, 1, ...

Certifique-se de que ***decresce*** a cada iteração. Caso contrário, reduza o valor de

Gradiente descendente**Gradiente descendente**

**Observações**

GD é sensível à escala dos dados

É recomendável aplicar normalização por padronização para fazer com que

GD é sensível à taxa de aprendizagem

Recomendável testar com ..., 0.001, 0.01, 0.1, 1, ...

Certifique-se de que ***decresce*** a cada iteração. Caso contrário, reduza o valor de

Tiago A. Almeida 

# iterações

**Ciência de dados**

**Regressão Linear Multivariada**

**Renato Moraes Silva**

**renato.silva@facens.br**

**Regressão linear multivariada**Tiago A. Almeida Regressão linear multivariada 

**Tamanho #Cômodos #Andares #Anos Valor'(U$1000)'**

2104 5 1% 45 460%

1416 3 2% 40 232%

1534% 3 2% 30 315%

852% 2% 1% 36% 178%

…% …% …% …% …%

**Notação:(**

="quan'dade"de"atributos

="""""""amostra"do"conjunto"de"treinamento

="valor"do"atributo""""da"""""""amostra"do"conjunto"de"

treinamento

Tiago A. Almeida

**Regressão linear multivariada** 

Regressão linear multivariada**Regressão linear univariada**

**Regressão linear multivariada**Tiago A. Almeida Regressão linear multivariada

**Regressão linear univariada**

**Regressão linear multivariada**

**Regressão linear multivariada**Tiago A. Almeida Regressão linear multivariada 

**Regressão linear univariada**

**Regressão linear multivariada**

Para simplificar:

**Regressão linear multivariada**Tiago A. Almeida Regressão linear multivariada 

**Regressão linear univariada**

**Regressão linear multivariada**

Para simplificar:

*h✓*(*x*) = *✓T x*

Tiago A. Almeida

**Regressão linear multivariada** 

Regressão linear multivariada

**Hipótese:)**

**Parâmetros:)**

**Função)de)custo:)**

**Gradiente descendente**

Repita

**Regressão linear multivariada**Tiago A. Almeida Regressão linear multivariada

**Algoritmo**

repetir até convergir {

*✓j* := *✓j ↵*1*m*X*m*

*i*=1

}

(*h✓*(*xi*) *yi*)*xij*

Atualizar “simultaneamente”

**Regressão linear multivariada**Tiago A. Almeida Regressão linear multivariada

*✓j* := *✓j ↵* 1*m*X*m i*=1

(*h✓*(*xi*) *yi*)*xij*

*✓*0 := *✓*0 *↵* 1*m*X*m i*=1

*✓*1 := *✓*1 *↵* 1*m*X*m i*=1

*✓*2 := *✓*2 *↵* 1*m*X*m i*=1

...

(*h✓*(*xi*) *yi*)*xi*0 (*h✓*(*xi*) *yi*)*xi*1 (*h✓*(*xi*) *yi*)*xi*2

**Medidas de desempenho**

*Mean squared error* (MSE):1*n*X*n t*=1

penaliza os erros maiores

*Mean absolute error* (MAE):1*n*X*n t*=1

*e*2*t*

*|et|*

*R*2(*coefficient of determination*): 1 *−SSres*

*SSres* =X*n*

*e*2*t* =X*n*

*SStot*, onde

*t*=1

*t*=1

(*yi − y*ˆ*i*)2, *SStot* =X*n t*=1

(*yi − y*¯)2, *yi* é o valor

verdadeiro, *y*ˆ*i* é o valor predito e *y*¯ =1*n*P*nt*=1*yi*

**melhor valor é 1

um modelo baseline que sempre prediz *y*¯ (média), independente dos atributos, possui valor 0.

pode assumir valores negativos (predições piores que a do método baseline)

**Exercício **

Calcule as métricas de desempenho MSE, MAE e *R*2, considerando a seguinte tabela de resultados esperados e resultados calculados pela regressão linear

Valor real Valor calculado

180 150

70 65

310 330

50 72

**Referências **

Aulas do Prof. Tiago Agostinho de Almeida (UFSCar, campus de Sorocaba)

Aulas do Prof. Juvenal J. Duarte

CARVALHO, André Carlos Ponce de Leon Ferreira et al. Inteligência Artificial - Uma Abordagem de Aprendizado de Máquina. Disponível em: Minha Biblioteca, (2nd edição). Grupo GEN, 2021.