



# **INTELIGÊNCIA ARTIFICIAL**

## **Uma Abordagem de Aprendizado de Máquina**

**2ª edição**

**Katti Faceli**

Professora-Associada da Universidade Federal de São Carlos (UFSCar)

**Ana Carolina Lorena**

Professora-Associada do Instituto Tecnológico de Aeronáutica (ITA)

**João Gama**

Professor Catedrático com Agregação da Universidade do Porto (Portugal)

**Tiago Agostinho de Almeida**

Professor-Associado da Universidade Federal de São Carlos (UFSCar)

**André Carlos Ponce de Leon Ferreira de Carvalho**

Professor Titular da Universidade de São Paulo (USP)

- ※ Os autores deste livro e a editora empenharam seus melhores esforços para assegurar que as informações e os procedimentos apresentados no texto estejam em acordo com os padrões aceitos à época da publicação, e todos os dados foram atualizados pelos autores até a data de fechamento do livro. Entretanto, tendo em conta a evolução das ciências, as atualizações legislativas, as mudanças regulamentares governamentais e o constante fluxo de novas informações sobre os temas que constam do livro, recomendamos enfaticamente que os leitores consultem sempre outras fontes fidedignas, de modo a se certificarem de que as informações contidas no texto estão corretas e de que não houve alterações nas recomendações ou na legislação regulamentadora.
- ※ Data do fechamento do livro: 04/02/2021
- ※ Os autores e a editora se empenharam para citar adequadamente e dar o devido crédito a todos os detentores de direitos autorais de qualquer material utilizado neste livro, dispondo-se a possíveis acertos posteriores caso, inadvertida e involuntariamente, a identificação de algum deles tenha sido omitida.
- ※ **Atendimento ao cliente: (11) 5080-0751 | faleconosco@grupogen.com.br**
- ※ Direitos exclusivos para a língua portuguesa  
Copyright © 2021 by  
**LTC | Livros Técnicos e Científicos Editora Ltda.**  
*Uma editora integrante do GEN | Grupo Editorial Nacional*  
Travessa do Ouvidor, 11  
Rio de Janeiro – RJ – 20040-040  
www.grupogen.com.br
- ※ Reservados todos os direitos. É proibida a duplicação ou reprodução deste volume, no todo ou em parte, em quaisquer formas ou por quaisquer meios (eletrônico, mecânico, gravação, fotocópia, distribuição pela Internet ou outros), sem permissão, por escrito, da LTC | Livros Técnicos e Científicos Editora Ltda.
- ※ Capa: Leônidas Leite
- ※ Imagem de capa: © PhonlamaiPhoto | iStockphoto.com
- ※ Edição eletrônica: 3Pontos Apoio Editorial Ltda.
- ※ Ficha catalográfica

**CIP-BRASIL. CATALOGAÇÃO NA PUBLICAÇÃO  
SINDICATO NACIONAL DOS EDITORES DE LIVROS, RJ**

---

148

2. ed.

Inteligência artificial : uma abordagem de aprendizado de máquina / Katti Faceli ... [et al.] . - 2. ed. - Rio de Janeiro : LTC, 2021.

: il. ; 28 cm.

Inclui bibliografia e índice

ISBN 978-85-216-3734-9

1. Inteligência artificial. 2. Inteligência artificial - Problemas, questões, exercícios. 3. Aprendizado do computador. I. Faceli, Katti.

20-68022

CDD: 006.3

CDU: 004.8

Leandra Felix da Cruz Candido - Bibliotecária - CRB-7/6135

---



# SUMÁRIO

<b>CAPÍTULO 1</b>	<b>INTRODUÇÃO.....</b>	<b>1</b>
1.1	Tarefas de Aprendizado.....	3
1.2	Aprendizado de Máquina e Indução de Modelos.....	4
1.3	Viés Indutivo .....	5
1.4	Estrutura do Livro.....	6
1.5	Exercícios .....	6

## **PARTE 1**      **PREPARAÇÃO DE DADOS**      **9**

<b>CAPÍTULO 2</b>	<b>ANÁLISE DE DADOS.....</b>	<b>12</b>
2.1	Caracterização de Dados .....	12
2.1.1	Tipo .....	13
2.1.2	Escala .....	14
2.2	Exploração de Dados.....	16
2.2.1	Dados Univariados.....	16
2.2.2	Dados Multivariados.....	22
2.3	Considerações Finais.....	26
2.4	Exercícios .....	26

<b>CAPÍTULO 3</b>	<b>PRÉ-PROCESSAMENTO DE DADOS.....</b>	<b>28</b>
3.1	Integração de Dados.....	29
3.2	Eliminação Manual de Atributos.....	29
3.3	Amostragem de Dados.....	30
3.4	Dados Desbalanceados.....	31
3.5	Limpeza de Dados .....	32
3.5.1	Dados Incompletos .....	32
3.5.2	Dados Inconsistentes .....	34
3.5.3	Dados Redundantes .....	35
3.5.4	Dados com Ruídos .....	37
3.6	Transformação de Dados .....	38
3.6.1	Conversão Simbólico-Numérico.....	38
3.6.2	Conversão Numérico-Simbólico.....	40
3.6.3	Transformação de Atributos Numéricos.....	41

3.7	Redução de Dimensionalidade .....	42
3.7.1	Agregação .....	43
3.7.2	Seleção de Atributos.....	43
3.7.3	Técnicas de Ordenação.....	44
3.7.4	Técnicas de Seleção de Subconjunto.....	45
3.8	Considerações Finais.....	46
3.9	Exercícios .....	46

## PARTE 2      MODELOS PREDITIVOS

49

<b>CAPÍTULO 4</b>	<b>MÉTODOS BASEADOS EM DISTÂNCIAS .....</b>	<b>53</b>
4.1	O Algoritmo do 1-Vizinho Mais Próximo.....	54
4.1.1	Superfícies de Decisão .....	55
4.1.2	Distâncias .....	55
4.2	O Algoritmo $k$ -NN.....	56
4.3	Análise do Algoritmo .....	57
4.3.1	Aspectos Positivos.....	57
4.3.2	Aspectos Negativos .....	57
4.4	Desenvolvimentos.....	58
4.5	Raciocínio Baseado em Casos.....	60
4.5.1	Representação de Casos.....	60
4.5.2	O Ciclo de Raciocínio Baseado em Casos.....	61
4.6	Considerações Finais.....	62
4.7	Exercícios .....	62
<b>CAPÍTULO 5</b>	<b>MÉTODOS PROBABILÍSTICOS .....</b>	<b>64</b>
5.1	Aprendizado Bayesiano.....	65
5.2	Classificador <i>Naive Bayes</i> .....	67
5.2.1	Detalhes de Implementação.....	68
5.2.2	Um Exemplo Ilustrativo .....	68
5.2.3	Análise do Algoritmo .....	71
5.2.4	Desenvolvimentos.....	71
5.3	Redes Bayesianas para Classificação.....	72
5.4	Considerações Finais.....	75
5.5	Exercícios .....	75
<b>CAPÍTULO 6</b>	<b>MÉTODOS SIMBÓLICOS .....</b>	<b>78</b>
6.1	Árvores de Decisão e Regressão.....	78
6.2	Indução de Árvores de Decisão e Regressão.....	80
6.2.1	Regras de Divisão para Classificação.....	80
6.2.2	Regras de Divisão para Regressão.....	85
6.2.3	Valores Desconhecidos.....	86
6.3	Estratégias de Poda .....	86
6.3.1	Pré-poda.....	87

6.3.2	Pós-poda.....	87
6.4	Análise do Algoritmo.....	88
6.4.1	Aspectos Positivos.....	89
6.4.2	Aspectos Negativos.....	89
6.5	Regras de Decisão.....	90
6.5.1	Por que Regras de Decisão?.....	90
6.5.2	De Árvores de Decisão às Regras de Decisão.....	91
6.5.3	O Algoritmo da Cobertura.....	91
6.6	Outros Modelos para Árvores de Previsão.....	96
6.6.1	Árvores de Modelos.....	96
6.6.2	Árvores de Opção.....	97
6.7	Considerações Finais.....	97
6.8	Exercícios.....	97
<b>CAPÍTULO 7</b>	<b>MÉTODOS CONEXIONISTAS.....</b>	<b>101</b>
7.1	Redes Neurais Artificiais.....	101
7.1.1	Sistema Nervoso.....	102
7.1.2	Componentes Básicos das RNAs.....	103
7.2	Redes Perceptron e Adaline.....	106
7.3	Perceptron Multicamadas.....	108
7.3.1	Algoritmo <i>Back-propagation</i> .....	109
7.3.2	Projeto da Arquitetura de uma RNA.....	112
7.4	Redes Profundas e Aprendizado Profundo.....	113
7.5	Análise do Algoritmo.....	114
7.5.1	Aspectos Positivos.....	115
7.5.2	Aspectos Negativos.....	115
7.6	Considerações Finais.....	115
7.7	Exercícios.....	116
<b>CAPÍTULO 8</b>	<b>MÉTODOS DE MAXIMIZAÇÃO DE MARGENS.....</b>	<b>117</b>
8.1	Teoria de Aprendizado Estatístico.....	117
8.1.1	Considerações sobre a Escolha do Classificador.....	118
8.1.2	Limites no Risco Esperado.....	118
8.2	SVMs Lineares.....	120
8.2.1	SVMs com Margens Rígidas.....	120
8.2.2	SVMs com Margens Suaves.....	122
8.3	SVMs Não Lineares.....	124
8.4	SVMs em Problemas de Regressão.....	127
8.5	Análise do Algoritmo.....	128
8.5.1	Aspectos Positivos.....	128
8.5.2	Aspectos Negativos.....	129
8.6	Considerações Finais.....	129
8.7	Exercícios.....	129

<b>CAPÍTULO 9</b>	<b>MODELOS MÚLTIPLOS PREDITIVOS .....</b>	<b>130</b>
9.1	Combinando Previsões de Classificadores .....	132
9.1.1	Métodos de Votação <i>versus</i> Métodos de Seriação.....	132
9.1.2	Métodos Dinâmicos <i>versus</i> Métodos Estáticos .....	133
9.2	Combinando Classificadores Homogêneos .....	135
9.2.1	Métodos Baseados em Amostragem dos Exemplos de Treinamento.....	136
9.2.2	Métodos Baseados na Amostragem do Conjunto de Atributos.....	139
9.2.3	Métodos Baseados na Injeção de Aleatoriedade .....	140
9.2.4	Métodos Baseados na Perturbação dos Exemplos de Teste.....	140
9.3	Combinando Classificadores Heterogêneos.....	141
9.3.1	Generalização em Pilha .....	141
9.3.2	Generalização em Cascata.....	142
9.3.3	Meta-Aprendizado .....	145
9.3.4	Sistemas Híbridos.....	145
9.4	Considerações Finais.....	146
9.5	Exercícios .....	146
<b>CAPÍTULO 10</b>	<b>AVALIAÇÃO DE MODELOS PREDITIVOS .....</b>	<b>148</b>
10.1	Métricas de Erro.....	149
10.1.1	Métricas para Classificação.....	149
10.1.2	Métricas para Regressão.....	150
10.2	Amostragem .....	150
10.2.1	<i>Holdout</i> e Amostragem Aleatória.....	151
10.2.2	Validação Cruzada.....	151
10.2.3	<i>Bootstrap</i> .....	152
10.3	Problemas de Duas Classes e o Espaço ROC .....	152
10.3.1	Medidas de Desempenho.....	153
10.3.2	Análise ROC.....	154
10.4	Testes de Hipóteses .....	156
10.4.1	Comparando Dois Modelos .....	157
10.4.2	Comparando Mais Modelos .....	158
10.5	Decomposição Viés-Variância da Taxa de Erro.....	160
10.5.1	Definição de Viés e Variância.....	160
10.5.2	Medindo os Componentes Viés-Variância.....	161
10.6	Considerações Finais.....	162
10.7	Exercícios .....	163
<b>PARTE 3</b>	<b>MODELOS DESCRITIVOS</b>	<b>165</b>
<b>CAPÍTULO 11</b>	<b>MINERAÇÃO DE PADRÕES FREQUENTES .....</b>	<b>167</b>
11.1	Mineração de Conjuntos de Itens Frequentes.....	167
11.1.1	O Espaço de Busca.....	168
11.2	O Algoritmo Apriori.....	169
11.2.1	Regras de Associação.....	170
11.2.2	Discussão .....	172

11.3	O Algoritmo FP-growth.....	172
11.4	Sumarização de <i>Itemsets</i> .....	174
11.4.1	Heurísticas para Seleção de Regras de Associação.....	174
11.5	Considerações Finais.....	176
11.6	Exercícios.....	176
<b>CAPÍTULO 12</b>	<b>ANÁLISE DE AGRUPAMENTOS.....</b>	<b>177</b>
12.1	Definições Básicas.....	178
12.2	Etapas da Análise de Agrupamento.....	181
12.2.1	Preparação dos Dados.....	182
12.2.2	Proximidade.....	183
12.2.3	Agrupamento.....	186
12.2.4	Validação.....	187
12.2.5	Interpretação.....	190
12.3	Considerações Finais.....	190
12.4	Exercícios.....	190
<b>CAPÍTULO 13</b>	<b>ALGORITMOS DE AGRUPAMENTO.....</b>	<b>192</b>
13.1	Algoritmos Hierárquicos.....	193
13.2	Algoritmos Particionais Baseados em Erro Quadrático.....	196
13.3	Algoritmos Baseados em Densidade.....	197
13.4	Algoritmos Baseados em Grafo.....	198
13.5	Algoritmos Baseados em Redes Neurais.....	198
13.6	Algoritmos Baseados em <i>Grid</i> .....	199
13.7	Considerações Finais.....	199
13.8	Exercícios.....	200
<b>CAPÍTULO 14</b>	<b>MODELOS MÚLTIPLOS DESCRITIVOS.....</b>	<b>202</b>
14.1	Comitês de Agrupamentos.....	204
14.1.1	Geração dos Agrupamentos Iniciais.....	205
14.1.2	Determinação da Função Consenso.....	206
14.1.3	Técnicas Baseadas em Comitês.....	208
14.2	Agrupamento com AEs Multiobjetivo.....	214
14.2.1	Técnicas Baseadas em AEs Multiobjetivo.....	216
14.3	Outras Técnicas.....	218
14.4	Considerações Finais.....	219
14.5	Exercícios.....	220
<b>CAPÍTULO 15</b>	<b>AValiação DE MODELOS DESCRITIVOS.....</b>	<b>221</b>
15.1	Critérios de Validação.....	222
15.2	Critérios Relativos.....	226
15.2.1	Índices Empregados em Critérios Relativos.....	226
15.2.2	Outras Abordagens de Validação Relativa.....	228
15.3	Critérios Internos.....	231
15.3.1	Estatística <i>Gap</i> .....	231

15.4	Critérios Externos.....	233
15.4.1	Índice Rand.....	234
15.4.2	Índice Jaccard .....	234
15.4.3	Índice de Fowlkes e Mallows.....	234
15.4.4	Índice Hubert normalizado .....	234
15.4.5	Índice Rand Corrigido .....	235
15.4.6	Índice Variação de Informação .....	235
15.4.7	Comparação dos Índices para Validação Externa .....	235
15.5	Considerações Finais.....	236
15.6	Exercícios .....	236

**PARTE 4 TÓPICOS ESPECIAIS (DISPONÍVEIS ONLINE) 239**

CAPÍTULO 16	APRENDIZADO EM FLUXOS CONTÍNUOS DE DADOS .....	241
CAPÍTULO 17	APRENDIZADO DE MÁQUINA AUTOMATIZADO .....	249
CAPÍTULO 18	DECOMPOSIÇÃO DE PROBLEMAS MULTICLASSE.....	256
CAPÍTULO 19	CLASSIFICAÇÃO MULTIRRÓTULO .....	264
CAPÍTULO 20	CLASSIFICAÇÃO HIERÁRQUICA.....	271
CAPÍTULO 21	COMPUTAÇÃO NATURAL .....	278
CAPÍTULO 22	ANÁLISE DE REDES SOCIAIS.....	285
CAPÍTULO 23	CATEGORIZAÇÃO DE TEXTOS.....	307

**PARTE 5 APLICAÇÕES (DISPONÍVEIS ONLINE) 321**

CAPÍTULO 24	AGRONEGÓCIOS .....	324
CAPÍTULO 25	ANÁLISE DE SENTIMENTO .....	326
CAPÍTULO 26	BIOINFORMÁTICA.....	329
CAPÍTULO 27	ECOLOGIA E MEIO AMBIENTE.....	332

<b>CAPÍTULO 28 ENERGIA</b> .....	<b>335</b>
<b>CAPÍTULO 29 FILTRAGEM DE MENSAGENS INDESEJADAS</b> .....	<b>338</b>
<b>CAPÍTULO 30 FINANÇAS</b> .....	<b>343</b>
<b>CAPÍTULO 31 MINERAÇÃO E CIÊNCIA DE DADOS</b> .....	<b>345</b>
<b>CAPÍTULO 32 ROBÓTICA</b> .....	<b>347</b>
<b>CAPÍTULO 33 SAÚDE</b> .....	<b>349</b>
<b>CAPÍTULO 34 SISTEMAS DE PERGUNTAS E RESPOSTAS</b> .....	<b>351</b>
<b>CAPÍTULO 35 SISTEMAS DE RECOMENDAÇÃO</b> .....	<b>354</b>
<b>PARTE 6 TENDÊNCIAS E PERSPECTIVAS</b> .....	<b>357</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	<b>363</b>
<b>ÍNDICE ALFABÉTICO</b> .....	<b>395</b>