

УДК 004.4  
ББК 32.972  
А79

**Араки М., Ватари М.**

А79 Занимательная манга. Машинное обучение: манга / Араки Масахиро (автор), Ватари Макана (худ.); пер. с яп. А. С. Слащевой. — М.: ДМК Пресс, 2020. — 214 с. : ил. — (Серия «Образовательная манга»). — Доп. тит. л. яп.

ISBN 978-5-97060-830-2

Сотруднику городской администрации Киёхара Кадзума поручено задание, которое без машинного обучения не выполнить. Под руководством своей давней знакомой Мияно Саяка он осваивает премудрости работы с искусственным интеллектом – от самых азов до глубокого обучения.

Вместе с героями манги читатели узнают о том, что такое регрессия и как проводить классификацию, ознакомятся с принципами оценки тестовых данных и особенностями работы нейронных сетей. В заключительной части излагаются методы обучения без учителя.

Издание предназначено для тех, кто начинает знакомство с машинным обучением и освоил математику на уровне первых курсов университета.

УДК 004.4  
ББК 32.972

Manga de Wakaru Kikai Gakushu (Manga Guide: Machine Learning )

By Araki Masaxiro (Author), Illustration by Vatari Makana

Published by Ohmsha, Ltd.

Russian language edition copyright © 2020 by DMK Press

Все права защищены. Никакая часть этого издания не может быть воспроизведена в любой форме или любыми средствами, электронными или механическими, включая фотографирование, ксерокопирование или иные средства копирования или сохранения информации, без письменного разрешения издательства.

ISBN 978-4-274-22244-3 (яп.)  
ISBN 978-5-97060-830-2 (рус.)

Copyright © 2018 by and Office sawa, Ltd.  
© Издание, перевод, ДМК Пресс, 2019

# СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ .....	V
-------------------	---

## Пролог

ПОГОВОРИМ О МАШИННОМ ОБУЧЕНИИ .....	1
-------------------------------------	---

В кабинете у Саяка (1). Саяка и старшеклассница Ай .....	14
--	----

## Глава 1

ЧТО ТАКОЕ РЕГРЕССИЯ .....	15
---------------------------	----

1.1. Сложности с прогнозом .....	16
----------------------------------	----

1.2. Определяем зависимые и независимые переменные .....	17
--	----

1.3. Находим функцию линейной регрессии .....	20
---	----

1.4. Регуляризация результата .....	22
-------------------------------------	----

В кабинете у Саяка (2). Математическое повторение (1) .....	34
---	----

## Глава 2

КАК ДЕЛАТЬ КЛАССИФИКАЦИЮ? .....	39
---------------------------------	----

2.1. Приводим данные в порядок .....	46
--------------------------------------	----

2.2. Определяем класс данных .....	47
------------------------------------	----

2.3. Логистическая регрессия .....	49
------------------------------------	----

2.4. Классификация по решающему дереву .....	55
--	----

В кабинете у Саяка (3). Математическое повторение (2) .....	74
---	----

## Глава 3

### ОЦЕНКА РЕЗУЛЬТАТОВ ..... 77

3.1. Без проверки тестовых данных никак нельзя.....	82
3.2. Обучающая, тестовая и проверочная выборки .....	83
3.3. Метод перекрестной проверки (кросс-валидации) .....	85
3.4. Доля правильно предсказанных объектов, точность, полнота и F-мера.....	87
В кабинете у Саяка (4). Математическое повторение (3).....	95

## Глава 4

### ГЛУБОКОЕ ОБУЧЕНИЕ ..... 97

4.1. Нейронная сеть .....	103
4.2. Обучение методом обратного распространения ошибок .....	107
4.3. Вызовы глубокого обучения.....	111
4.3.1. Проблема глубокой нейронной сети .....	112
4.3.2. Хитрости многоступенчатого обучения	
1. Метод предварительного обучения .....	113
4.3.3. Хитрости многоступенчатого обучения	
2. Функция активации.....	115
4.3.4. Хитрости многоступенчатого обучения	
3. Как избежать переобучения.....	117
4.3.5. Нейронные сети со специализированной структурой.....	118
В кабинете у Саяка (5). Математическое повторение (4) .....	134

## Глава 5

### АНСАМБЛЕВЫЕ МЕТОДЫ ..... 139

5.1. Бэггинг .....	146
5.2. Случайный лес.....	149
5.3. Бустинг .....	152
В кабинете у Саяка (6). Математическое повторение (5) .....	160

<b>Глава 6</b>	
<b>ОБУЧЕНИЕ БЕЗ УЧИТЕЛЯ .....</b>	<b>165</b>
6.1. Кластеризация.....	172
6.1.1. Иерархическая кластеризация .....	173
6.1.2. Разделяющая кластеризация.....	175
6.2. Разложение матрицы.....	179
В кабинете у Саяка (7). Математическое повторение (6) .....	191
<b>ЭПИЛОГ .....</b>	<b>197</b>
<b>ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ .....</b>	<b>205</b>