

ББК 32.973.26-018.2.75

Ч-49

УДК 681.3.07

ООО “Диалектика”

Зав. редакцией *С.Н. Тригуб*

Перевод с английского и редакция *В.А. Коваленко*

По общим вопросам обращайтесь в издательство “Диалектика” по адресу:
info@dialektika.com, <http://www.dialektika.com>

Черняк, Евгений.

Ч-49 Введение в глубокое обучение. : Пер. с англ. — СПб. : ООО “Диалектика”, 2020. — 192 с. : ил. — Парал. тит. англ.

ISBN 978-5-907203-10-5 (рус.)

ББК 32.973.26-018.2.75

Все названия программных продуктов являются зарегистрированными торговыми марками соответствующих фирм.

Никакая часть настоящего издания ни в каких целях не может быть воспроизведена в какой бы то ни было форме и какими бы то ни было средствами, будь то электронные или механические, включая фотокопирование и запись на магнитный носитель, если на это нет письменного разрешения издательства MIT Press.

Copyright © 2020 by Dialektika Computer Publishing.

Authorized translation from the English language edition *Introduction to Deep Learning* (ISBN 978-0-262-03951-2) published by MIT Press, Copyright © 2018 by Massachusetts Institute of Technology.

All rights reserved. No part of this book may be reproduced in any form or by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

Научно-популярное издание

Евгений Черняк

Введение в глубокое обучение

ООО “Диалектика”, 195027, Санкт-Петербург, Магнитогорская ул., д. 30, лит. А, пом. 848

ISBN 978-5-907203-10-5 (рус.)

© ООО “Диалектика”, 2020,
перевод, оформление, макетирование

ISBN 978-0-262-03951-2 (англ.)

© 2018 The Massachusetts Institute of Technology

Оглавление

Введение	9
Глава 1. Нейронные сети с прямой связью	11
Глава 2. Язык Tensorflow	39
Глава 3. Сверточные нейронные сети	61
Глава 4. Векторное представление слов и рекуррентные NN	79
Глава 5. Обучение от последовательности к последовательности	103
Глава 6. Глубокое обучение с подкреплением	121
Глава 7. Модели нейронных сетей, обучаемых без учителя	145
Приложение. Выборочные ответы к упражнениям	169
Предметный указатель	175

Содержание

Введение	9
Посвящение	10
Ждем ваших отзывов!	10
Глава 1. Нейронные сети с прямой связью	11
1.1. Перцептроны	13
1.2. Функция кросс-энтропийных потерь для нейронных сетей	19
1.3. Производные и стохастический градиентный спуск	23
1.4. Написание программы	28
1.5. Матричное представление нейронных сетей	30
1.6. Независимость данных	33
1.7. Ссылки и что читать дальше	35
1.8. Упражнения	36
Глава 2. Язык Tensorflow	39
2.1. Вводная информация о Tensorflow	39
2.2. Программа TF	43
2.3. Многослойные NN	48
2.4. Другие части	51
2.4.1. Создание контрольных точек	51
2.4.2. <code>tf.nn</code>	53
2.4.3. Инициализация переменных TF	55
2.4.4. Упрощение создания графов TF	57
2.5. Ссылки и что читать дальше	58
2.6. Упражнения	58
Глава 3. Сверточные нейронные сети	61
3.1. Фильтры, шаги и дополнения	62
3.2. Простой пример свертки TF	68
3.3. Многослойная свертка	70
3.4. Детали свертки	74
3.4.1. Смещения	74
3.4.2. Слои со сверткой	74
3.4.3. Объединение	75

СОДЕРЖАНИЕ	7
3.5. Ссылки и что читать дальше	76
3.6. Упражнения	77
Глава 4. Векторное представление слов и рекуррентные NN	79
4.1. Векторное представление слова для языковых моделей	79
4.2. Создание языковых моделей с прямой связью	84
4.3. Улучшение языковых моделей с прямой связью	86
4.4. Переобучение	87
4.5. Рекуррентные сети	90
4.6. Долгая краткосрочная память	97
4.7. Ссылки и что читать дальше	100
4.8. Упражнения	101
Глава 5. Обучение от последовательности к последовательности	103
5.1. Парадигма Seq2Seq	104
5.2. Написание программы Seq2Seq MT	107
5.3. Внимание в Seq2seq	110
5.4. Seq2Seq с несколькими длинами окон	114
5.5. Упражнение по программированию	115
5.6. Упражнения	118
5.7. Ссылки и что читать дальше	118
Глава 6. Глубокое обучение с подкреплением	121
6.1. Итерация по значениям	122
6.2. Q-обучение	125
6.3. Основы глубокого Q-обучения	128
6.4. Методы градиента политики	131
6.5. Методы актер-критик	138
6.6. Воспроизведение опыта	140
6.7. Ссылки и что читать дальше	142
6.8. Упражнения	142
Глава 7. Модели нейронных сетей, обучаемых без учителя	145
7.1. Основы автокодировки	145
7.2. Сверточное автокодирование	149
7.3. Вариационное автокодирование	153
7.4. Генеративно-состязательные сети	161

7.5. Ссылки и что читать дальше	166
7.6. Упражнения	166
Приложение. Ответы к некоторым упражнениям	169
Глава 1	169
Глава 2	170
Глава 3	170
Глава 4	171
Глава 5	171
Глава 6	172
Глава 7	173
Предметный указатель	175