

Краткое содержание

ЧАСТЬ I. ОСНОВЫ ГЛУБОКОГО ОБУЧЕНИЯ	25
Глава 1. Что такое глубокое обучение?	26
Глава 2. Прежде чем начать: математические основы нейронных сетей . .	51
Глава 3. Начало работы с нейронными сетями	81
Глава 4. Основы машинного обучения	120
ЧАСТЬ II. ГЛУБОКОЕ ОБУЧЕНИЕ НА ПРАКТИКЕ	147
Глава 5. Глубокое обучение в технологиях компьютерного зрения	148
Глава 6. Глубокое обучение для текста и последовательностей	210
Глава 7. Лучшие практики глубокого обучения продвинутого уровня . . .	270
Глава 8. Генеративное глубокое обучение	307
Глава 9. Заключение	357
Приложение А. Установка Keras и его зависимостей в Ubuntu	384
Приложение В. Запуск Jupyter Notebook на экземпляре EC2 GPU	389

Оглавление

Предисловие	16
Благодарности	17
Об этой книге	19
Кому адресована эта книга	19
Содержание книги	20
Требования к программному/аппаратному обеспечению	21
Исходный код	22
Форум книги	22
От издательства	22
Об авторе	23
Об иллюстрации на обложке	24
ЧАСТЬ I. ОСНОВЫ ГЛУБОКОГО ОБУЧЕНИЯ	25
Глава 1. Что такое глубокое обучение?	26
1.1. Искусственный интеллект, машинное и глубокое обучение	27
1.1.1. Искусственный интеллект	27
1.1.2. Машинное обучение	28
1.1.3. Обучение представлению данных	29
1.1.4. «Глубина» глубокого обучения	31
1.1.5. Принцип действия глубокого обучения в трех картинках	33
1.1.6. Какой ступени развития достигло глубокое обучение	35

1.1.7. Не верьте рекламе	36
1.1.8. Перспективы ИИ	37
1.2. Что было до глубокого обучения: краткая история машинного обучения . . .	38
1.2.1. Вероятностное моделирование	38
1.2.2. Первые нейронные сети	39
1.2.3. Ядерные методы	39
1.2.4. Деревья решений, случайные леса и градиентный бустинг	41
1.2.5. Назад к нейронным сетям	42
1.2.6. Отличительные черты глубокого обучения	43
1.2.7. Современный ландшафт машинного обучения	44
1.3. Почему глубокое обучение? Почему сейчас?	45
1.3.1. Оборудование	45
1.3.2. Данные	46
1.3.3. Алгоритмы	47
1.3.4. Новая волна инвестиций	48
1.3.5. Демократизация глубокого обучения	48
1.3.6. Ждать ли продолжения этой тенденции?	49

Глава 2. Прежде чем начать: математические основы нейронных сетей 51

2.1. Первое знакомство с нейронной сетью	51
2.2. Представление данных для нейронных сетей	55
2.2.1. Скаляры (тензоры нулевого ранга)	56
2.2.2. Векторы (тензоры первого ранга)	56
2.2.3. Матрицы (тензоры второго ранга)	56
2.2.4. Тензоры третьего и высшего рангов	57
2.2.5. Ключевые атрибуты	57
2.2.6. Манипулирование тензорами с помощью NumPy	59
2.2.7. Пакеты данных	59
2.2.8. Практические примеры тензоров с данными	60
2.2.9. Векторные данные	60
2.2.10. Временные ряды или последовательности	61

2.2.11. Изображения	61
2.2.12. Видео	62
2.3. Шестеренки нейронных сетей: операции с тензорами	63
2.3.1. Поэлементные операции.	63
2.3.2. Расширение	64
2.3.3. Скалярное произведение тензоров	65
2.3.4. Изменение формы тензора	68
2.3.5. Геометрическая интерпретация операций с тензорами	69
2.3.6. Геометрическая интерпретация глубокого обучения	70
2.4. Механизм нейронных сетей: оптимизация на основе градиента	71
2.4.1. Что такое производная?	72
2.4.2. Производная операций с тензорами: градиент	73
2.4.3. Стохастический градиентный спуск.	74
2.4.4. Объединение производных: алгоритм обратного распространения ошибки.	77
2.5. Оглядываясь на первый пример	78
Краткие итоги главы	79

Глава 3. Начало работы с нейронными сетями 81

3.1. Анатомия нейронной сети	82
3.1.1. Слои: строительные блоки глубокого обучения.	83
3.1.2. Модели: сети слоев	84
3.1.3. Функции потерь и оптимизаторы: ключи к настройке процесса обучения	84
3.2. Введение в Keras	85
3.2.1. Keras, TensorFlow, Theano и CNTK	86
3.2.2. Разработка с использованием Keras: краткий обзор.	87
3.3. Настройка рабочей станции для глубокого обучения	89
3.3.1. Jupyter Notebook: предпочтительный способ проведения экспериментов с глубоким обучением	90
3.3.2. Подготовка Keras: два варианта	90
3.3.3. Запуск заданий глубокого обучения в облаке: за и против	91
3.3.4. Выбор GPU для глубокого обучения	91

3.4. Классификация отзывов к фильмам: пример бинарной классификации	92
3.4.1. The IMDB dataset	92
3.4.2. Подготовка данных	93
3.4.3. Конструирование сети	94
3.4.4. Проверка решения	98
3.4.5. Использование обученной сети для предсказаний на новых данных	101
3.4.6. Дальнейшие эксперименты	101
3.4.7. Подведение итогов	102
3.5. Классификация новостных лент: пример классификации в несколько классов	102
3.5.1. Набор данных Reuters	103
3.5.2. Подготовка данных	104
3.5.3. Конструирование сети	105
3.5.4. Проверка решения	106
3.5.5. Предсказания на новых данных	108
3.5.6. Другой способ обработки меток и потерь	109
3.5.7. Важность использования достаточно больших промежуточных слоев	109
3.5.8. Дальнейшие эксперименты	110
3.5.9. Подведение итогов	110
3.6. Предсказание цен на дома: пример регрессии	111
3.6.1. Набор данных с ценами на жилье в Бостоне	111
3.6.2. Подготовка данных	112
3.6.3. Конструирование сети	112
3.6.4. Оценка решения методом перекрестной проверки по K блокам	113
3.6.5. Подведение итогов	118
Краткие итоги главы	118

Глава 4. Основы машинного обучения 120

4.1. Четыре раздела машинного обучения	120
4.1.1. Контролируемое обучение	121
4.1.2. Неконтролируемое обучение	121

4.1.3. Самоконтролируемое обучение	122
4.1.4. Обучение с подкреплением	122
4.2. Оценка моделей машинного обучения	124
4.2.1. Тренировочные, проверочные и контрольные наборы данных	124
4.2.2. Что важно помнить	128
4.3. Обработка данных, конструирование признаков и обучение признаков	128
4.3.1. Предварительная обработка данных для нейронных сетей	129
4.3.2. Конструирование признаков	130
4.4. Переобучение и недообучение	132
4.4.1. Уменьшение размера сети	133
4.4.2. Добавление регуляризации весов	136
4.4.3. Добавление прореживания	138
4.5. Обобщенный процесс решения задач машинного обучения	140
4.5.1. Определение задачи и создание набора данных	140
4.5.2. Выбор меры успеха	141
4.5.3. Выбор протокола оценки	142
4.5.4. Предварительная подготовка данных	142
4.5.5. Разработка модели, более совершенной, чем базовый случай	143
4.5.6. Масштабирование по вертикали: разработка модели с переобучением	144
4.5.7. Регуляризация модели и настройка гиперпараметров	145
Краткие итоги главы	146

ЧАСТЬ II. ГЛУБОКОЕ ОБУЧЕНИЕ НА ПРАКТИКЕ 147

Глава 5. Глубокое обучение в технологиях компьютерного зрения 148

5.1. Введение в сверточные нейронные сети	148
5.1.1. Операция свертывания	151
5.1.2. Выбор максимального значения из соседних (max-pooling)	157
5.2. Обучение сверточной нейронной сети с нуля на небольшом наборе данных	159
5.2.1. Целесообразность глубокого обучения для решения задач с небольшими наборами данных	159

5.2.2. Загрузка данных	160
5.2.3. Конструирование сети	163
5.2.4. Предварительная обработка данных	165
5.2.5. Расширение данных	169
5.3. Использование предварительно обученной сверточной нейронной сети . . .	173
5.3.1. Выделение признаков	174
5.3.2. Дообучение	184
5.3.3. Подведение итогов	190
5.4. Визуализация знаний, заключенных в сверточной нейронной сети	191
5.4.1. Визуализация промежуточных активаций	191
5.4.2. Визуализация фильтров сверточных нейронных сетей	198
5.4.3. Визуализация тепловых карт активации класса	204
Краткие итоги главы	209

Глава 6. Глубокое обучение для текста и последовательностей . . 210

6.1. Работа с текстовыми данными	211
6.1.1. Прямое кодирование слов и символов	213
6.1.2. Векторное представление слов	215
6.1.3. Объединение всего вместе: от исходного текста к векторному представлению слов	221
6.1.4. Подведение итогов	228
6.2. Рекуррентные нейронные сети	228
6.2.1. Рекуррентный слой в Keras	231
6.2.2. Слои LSTM и GRU	235
6.2.3. Пример использования слоя LSTM из Keras	238
6.2.4. Подведение итогов	240
6.3. Улучшенные методы использования рекуррентных нейронных сетей	240
6.3.1. Задача прогнозирования температуры	241
6.3.2. Подготовка данных	244
6.3.3. Базовое решение без привлечения машинного обучения	247
6.3.4. Базовое решение с привлечением машинного обучения	248
6.3.5. Первое базовое рекуррентное решение	250

6.3.6. Использование рекуррентного прореживания для борьбы с переобучением	251
6.3.7. Наложение нескольких рекуррентных слоев друг на друга	253
6.3.8. Использование двунаправленных рекуррентных нейронных сетей	254
6.3.9. Что дальше	258
6.3.10. Подведение итогов	259
6.4. Обработка последовательностей с помощью сверточных нейронных сетей	260
6.4.1. Обработка последовательных данных с помощью одномерной сверточной нейронной сети	260
6.4.2. Выбор соседних значений в одномерной последовательности данных	261
6.4.3. Реализация одномерной сверточной сети	262
6.4.4. Объединение сверточных и рекуррентных сетей для обработки длинных последовательностей	264
6.4.5. Подведение итогов	268
Краткие итоги главы	268

Глава 7. Лучшие практики глубокого обучения продвинутого уровня 270

7.1. За рамками модели Sequential: функциональный API фреймворка Keras	270
7.1.1. Введение в функциональный API	274
7.1.2. Модели с несколькими входами.	275
7.1.3. Модели с несколькими выходами	277
7.1.4. Ориентированные ациклические графы уровней	280
7.1.5. Повторное использование экземпляров слоев	284
7.1.6. Модели как слои.	285
7.1.7. Подведение итогов	286
7.2. Исследование и мониторинг моделей глубокого обучения с использованием обратных вызовов Keras и TensorBoard	286
7.2.1. Применение обратных вызовов для воздействия на модель в ходе обучения	287
7.2.2. Введение в TensorBoard: фреймворк визуализации TensorFlow.	290
7.2.3. Подведение итогов	296

7.3. Извлечение максимальной пользы из моделей	297
7.3.1. Шаблоны улучшенных архитектур.	297
7.3.2. Оптимизация гиперпараметров.	301
7.3.3. Ансамблирование моделей	303
7.3.4. Подведение итогов	305
Краткие итоги главы	306

Глава 8. Генеративное глубокое обучение 307

8.1. Генерирование текста с помощью LSTM	309
8.1.1. Краткая история генеративных рекуррентных сетей	309
8.1.2. Как генерируются последовательности данных?	310
8.1.3. Важность стратегии выбора	311
8.1.4. Реализация посимвольной генерации текста на основе LSTM	313
8.1.5. Подведение итогов	318
8.2. DeepDream	318
8.2.1. Реализация DeepDream в Keras.	320
8.2.2. Подведение итогов	326
8.3. Нейронная передача стиля	326
8.3.1. Функция потерь содержимого.	327
8.3.2. Функция потерь стиля	328
8.3.3. Нейронная передача стиля в Keras	329
8.3.4. Подведение итогов	336
8.4. Генерирование изображений с вариационными автокодировщиками	336
8.4.1. Выбор шаблонов из скрытых пространств изображений	336
8.4.2. Концептуальные векторы для редактирования изображений.	338
8.4.3. Вариационные автокодировщики	339
8.4.4. Подведение итогов	346
8.5. Введение в генеративно-состязательные сети	346
8.5.1. Реализация простейшей генеративно-состязательной сети	348
8.5.2. Набор хитростей.	349
8.5.3. Генератор	350
8.5.4. Дискриминатор.	351

8.5.5. Состязательная сеть	352
8.5.6. Как обучить сеть DCGAN	353
8.5.7. Подведение итогов	355
Краткие итоги главы	356

Глава 9. Заключение 357

9.1. Краткий обзор ключевых понятий	358
9.1.1. Разные подходы к ИИ	358
9.1.2. Что делает глубокое обучение особенным среди других подходов к машинному обучению	358
9.1.3. Как правильно воспринимать глубокое обучение	359
9.1.4. Ключевые технологии.	360
9.1.5. Обобщенный процесс машинного обучения.	361
9.1.6. Основные архитектуры сетей	362
9.1.7. Пространство возможностей	367
9.2. Ограничения глубокого обучения	369
9.2.1. Риск очеловечивания моделей глубокого обучения.	370
9.2.2. Локальное и экстремальное обобщение	372
9.2.3. Подведение итогов.	373
9.3. Будущее глубокого обучения	374
9.3.1. Модели как программы.	375
9.3.2. За границами алгоритма обратного распространения ошибки и дифференцируемых слоев	377
9.3.3. Автоматизированное машинное обучение	377
9.3.4. Непрерывное обучение и повторное использование модульных подпрограмм	378
9.3.5. Долгосрочная перспектива	380
9.4. Как не отстать от прогресса в быстроразвивающейся области	381
9.4.1. Практические решения реальных задач на сайте Kaggle	381
9.4.2. Знакомство с последними разработками на сайте arXiv	382
9.4.3. Исследование экосистемы Keras	383
9.5. Заключительное слово	383

Приложение А. Установка Keras и его зависимостей в Ubuntu . . . 384

А.1. Установка пакетов научных вычислений для Python	385
А.2. Настройка поддержки GPU	386
А.3. Установка Theano (необязательно)	387
А.4. Установка Keras	388

Приложение В. Запуск Jupyter Notebook на экземпляре EC2 GPU 389

В.1. Что такое Jupyter Notebook? Зачем запускать Jupyter Notebook на AWS GPU?	389
В.2. Когда нежелательно использовать Jupyter на AWS для глубокого обучения?	390
В.3. Настройка экземпляра AWS GPU	390
В.3.1. Настройка Jupyter.	393
В.4. Установка Keras	394
В.5. Настройка перенаправления локальных портов	395
В.6. Доступ к Jupyter из браузера на локальном компьютере	395