

# Содержание

---

Об авторе	13
О рецензенте	15
<b>Введение</b>	<b>17</b>
Для кого предназначена эта книга	18
Структура книги	18
Файлы примеров	20
Ждем ваших отзывов!	21
<b>Глава 1</b>	
<b>Математические основы</b>	<b>23</b>
Линейная алгебра	24
Векторы	24
Скаляры	26
Матрицы	26
Тензоры	26
Операции над матрицами	27
Линейная независимость векторов	31
Ранг матрицы	33
Единичная матрица	34
Определитель матрицы	34
Обратная матрица	36
Норма вектора	38

Псевдообращение матрицы	39
Единичный вектор в направлении заданного вектора	40
Проекция вектора на направление другого вектора	40
Собственные векторы	40
Математический анализ	47
Дифференцирование	48
Градиент функции	49
Смешанные частные производные	49
Гессиан функции	50
Экстремумы функций	50
Локальные и глобальные минимумы	52
Положительно полуопределенные и положительно определенные матрицы	54
Выпуклые множества	54
Выпуклые функции	54
Невыпуклые функции	55
Примеры выпуклых и невыпуклых функций нескольких переменных	56
Ряд Тейлора	58
Теория вероятностей	59
Объединение, пересечение и условная вероятность событий	60
Цепное правило для вычисления вероятности пересечения событий	62
Взаимоисключающие события	62
Независимость событий	62
Условная независимость событий	62
Теорема Байеса	63
Функция вероятности	63
Плотность вероятности	63
Математическое ожидание случайной переменной	64
Дисперсия случайной переменной	64
Асимметрия и эксцесс	65
Ковариация	68
Коэффициент корреляции	68
Распространенные типы распределений	69
Функция правдоподобия	75
Метод максимального правдоподобия	76
Проверка статистических гипотез и $p$ -значение	77
Алгоритмы машинного обучения и методы оптимизации	80
Обучение с учителем	80
Обучение без учителя	91
Методы оптимизации, используемые в машинном обучении	91
Оптимизация при наличии ограничений	103

## 8 СОДЕРЖАНИЕ

Некоторые важные темы машинного обучения	105
Методы снижения размерности данных	105
Регуляризация	110
Регуляризация, рассматриваемая как задача оптимизации с ограничениями	113
Резюме	114

## Глава 2

### **Введение в глубокое обучение и TensorFlow** 115

Эволюция глубокого обучения	115
Перцептроны и алгоритм обучения перцептрона	118
Геометрическая интерпретация процесса обучения перцептрона	122
Ограничения процесса обучения перцептронов	124
Потребность в нелинейности	127
Функция активации перцептронов для скрытого слоя, обеспечивающая нелинейность	128
Типы функций активации для нейрона/перцептрона	130
Правила обучения для многослойной перцептронной сети	136
Вычисление градиента с помощью метода обратного распространения ошибки	138
Вычисление градиента с помощью обобщенного метода обратного распространения ошибки	139
Глубокое обучение и традиционные методы	145
Библиотека TensorFlow	146
Пакеты глубокого обучения общего назначения	147
Установка TensorFlow	149
Основы разработки с помощью TensorFlow	149
Сравнение методов оптимизации на основе градиентного спуска с точки зрения глубокого обучения	153
Скорость обучения при использовании мини-пакетов в методе стохастического градиентного спуска	160
Оптимизаторы TensorFlow	160
Реализация логики XOR средствами TensorFlow	170
Линейная регрессия в TensorFlow	175
Многоклассовая классификация при помощи полнопакетного градиентного спуска с использованием функции активации SoftMax	179
Многоклассовая классификация при помощи стохастического градиентного спуска с использованием функции активации SoftMax	182
GPU	185
Резюме	185

**Глава 3****Сверточные нейронные сети**

	187
Операция свертки	187
Линейные стационарные системы	188
Свертка сигналов в одном измерении	190
Аналоговые и цифровые сигналы	192
2D- и 3D-сигналы	195
2D-свертка	195
Двумерная единичная дельта-функция	197
2D-свертка сигнала с откликом LSI-системы на единичный дельта-импульс	198
2D-свертка изображения с импульсным откликом различных LSI-систем	200
Часто используемые фильтры изображений	204
Усредняющий фильтр	205
Медианный фильтр	207
Гауссовский фильтр	209
Градиентные фильтры	209
Выделение краев с помощью фильтра Собеля	212
Тождественное преобразование	214
Сверточные нейронные сети	215
Компоненты сверточных нейронных сетей	215
Входной слой	217
Слой свертки	217
Слой субдискретизации	219
Обратное распространение ошибки через сверточный слой	220
Обратное распространение ошибки через слои субдискретизации	224
Разделение весов посредством свертки и его преимущества	225
Трансляционная эквивариантность	226
Трансляционная инвариантность при выполнении субдискретизации	228
Исключение слоев и регуляризация	229
Сверточная нейронная сеть для распознавания цифр из набора данных MNIST	231
Применение сверточных нейронных сетей для решения реальных задач	236
Пакетная нормализация	244
Варианты архитектур сверточных нейронных сетей	247
LeNet	248
AlexNet	249
VGG16	251

## 10 СОДЕРЖАНИЕ

ResNet	252
Переносимое обучение	253
Рекомендации по использованию переносимого обучения	255
Переносимое обучение с помощью сети InceptionV3 компании Google	256
Переносимое обучение с помощью предварительно обученной сети VGG16	260
Резюме	266

## Глава 4

### **Обработка естественного языка с использованием рекуррентных нейронных сетей**

	267
Векторная модель	268
Векторное представление слов	271
Технология Word2Vec	273
Метод CBOW	273
Реализация модели CBOW в TensorFlow	277
Модель векторного представления слов на основе сkip-грамм	280
Реализация модели на основе сkip-грамм в TensorFlow	284
Векторы слов, основанные на глобальных статистиках совместной встречаемости слов	288
Библиотека GloVe	294
Обработка словесных аналогий с помощью векторов слов	298
Введение в рекуррентные нейронные сети	301
Языковое моделирование	304
Предсказание следующего слова предложения традиционными методами и с помощью RNN	305
Алгоритм обратного распространения ошибки во времени (BPTT)	306
Проблема затухающих и взрывных градиентов в RNN	309
Решение проблемы затухающих и взрывных градиентов в RNN	311
Долгая краткосрочная память (LSTM)	313
Сглаживание проблемы взрывных и затухающих градиентов с помощью LSTM	315
Распознавание цифр набора MNIST средствами TensorFlow с использованием рекуррентных нейронных сетей	316
Управляемые рекуррентные блоки	327
Двунаправленные RNN	329
Резюме	331

**Глава 5****Обучение без учителя с использованием ограниченных машин Больцмана и автокодировщиков**

	333
Распределение Больцмана	334
Байесовский вывод: правдоподобие, априорные и апостериорные распределения вероятностей	335
Семплирование методами Монте-Карло по схеме марковских цепей	342
Алгоритм Метрополиса	346
Ограниченные машины Больцмана	351
Обучение ограниченной машины Больцмана	356
Семплирование по Гиббсу	361
Блочное семплирование по Гиббсу	363
Стадия приработки и генерирование выборок при семплировании по Гиббсу	364
Использование семплирования по Гиббсу в ограниченных машинах Больцмана	364
Контрастная дивергенция	366
Реализация ограниченной машины Больцмана средствами TensorFlow	367
Коллаборативная фильтрация с использованием ограниченных машин Больцмана	372
Глубокие сети доверия	377
Автокодировщики	383
Обучение признакам посредством автокодировщиков в целях обучения с учителем	385
Расхождение Кульбака — Лейблера	388
Разреженные автокодировщики	388
Реализация разреженного автокодировщика в TensorFlow	390
Шумоподавляющие автокодировщики	394
Реализация шумоподавляющего автокодировщика в TensorFlow	395
PCA- и ZCA-отбеливание	402
Резюме	406

**Глава 6****Усовершенствованные варианты архитектуры нейронных сетей**

	407
Сегментация изображений	408
Бинарный пороговый метод, основанный на гистограммах интенсивности пикселей	408

## 12 СОДЕРЖАНИЕ

Метод Оцу	409
Сегментация изображений методом водоразделов	413
Сегментация изображений методом кластеризации с помощью K-средних	416
Семантическая сегментация	419
Метод скользящего окна	419
Полносверточная сеть	420
Полносверточная сеть с понижающей и повышающей дискретизацией	423
U-Net	427
Семантическая сегментация с помощью полносвязных нейронных сетей средствами TensorFlow	431
Сети классификации и локализации изображений	441
Обнаружение объектов	444
Сеть R-CNN	445
Сети Fast и Faster R-CNN	447
Генеративно-сопоставительные сети	448
Задачи максимина и минимакса	450
Игра с нулевой суммой	452
Минимакс и седловые точки	454
Функция стоимости и тренировка сети GAN	456
Затухающие градиенты для генератора	458
Реализация сети GAN средствами TensorFlow	459
Развертывание моделей TensorFlow в производственной среде	463
Резюме	466
<b>Предметный указатель</b>	<b>467</b>