

ГЛУБОКОЕ ОБУЧЕНИЕ

ПОГРУЖЕНИЕ В МИР НЕЙРОННЫХ СЕТЕЙ



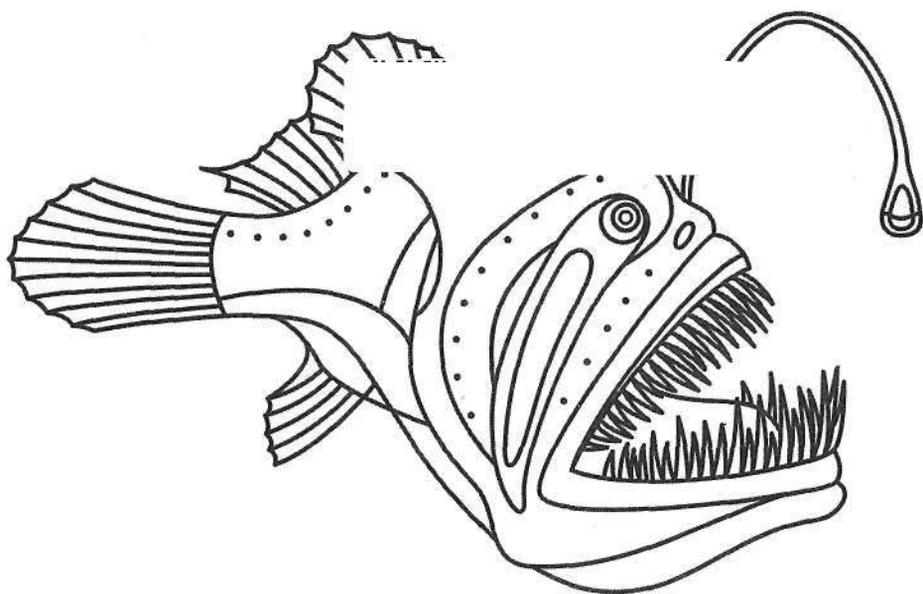
$$\begin{aligned}
 \text{KL}(q(w) \| p(w) | X, Y) &= \int q(w) \log \frac{q(w)}{p(w)} p(w) | X, Y \, dw = \\
 &= \int q(w) \log \frac{q(w)}{p(w)} p(Y | X, w) \, dw + \text{const} = \\
 &= - \int q(w) \log p(Y | X, w) \, dw + \int q(w) \log \frac{q(w)}{p(w)} \, dw + \text{const} = \\
 &= - \int q(w) \log p(Y | X, w) \, dw + \text{KL}(q(w) \| p(w)) + \text{const} = \\
 &= - \sum_{i=1}^N \int q(w) \log p(y_i | f_w(x_i)) \, dw + \text{KL}(q(w) \| p(w)) + \text{const},
 \end{aligned}$$



С. Николенко, А. Кадулин, Е. Архангельская

С. Николенко, А. Кадулин, Е. Архангельская

ГЛУБОКОЕ ОБУЧЕНИЕ ПОГРУЖЕНИЕ В МИР НЕЙРОННЫХ СЕТЕЙ



 ПИТЕР®

Санкт-Петербург • Москва • Екатеринбург • Воронеж
Нижний Новгород • Ростов-на-Дону • Самара • Минск

2019

Николенко, С. И. Глубокое обучение : погружение в мир нейронных сетей / С. И. Николенко, А. А. Кадулин, Е. О. Архангельская. — Санкт-Петербург [и др.] : Питер, 2019. — 476 с. : ил. — (Библиотека программиста).

УДК 004.8.032.26

ББК 32

Чит. зал №1 — 2 экз.

Перед вами — первая книга о глубоком обучении, написанная на русском языке. Глубокие модели оказались ключом, который подходит ко всем замкам сразу: новые архитектуры и алгоритмы обучения, а также увеличившиеся вычислительные мощности и появившиеся огромные наборы данных привели к революционным прорывам в компьютерном зрении, распознавании речи, обработке естественного языка и многих других типично «человеческих» задачах машинного обучения. Эти захватывающие идеи, вся история и основные компоненты революции глубокого обучения, а также самые современные достижения этой области доступно и интересно изложены в книге. Максимум объяснений, минимум кода, серьезный материал о машинном обучении и увлекательное изложение — в этой уникальной работе замечательных российских ученых и интеллектуалов.

Оглавление

Часть I. Как обучать нейронные сети

Глава 1. От биологии к информатике, или We need to go deeper	6
1.1. Революция обучения глубоких сетей	7
1.2. Искусственный интеллект и машинное обучение	11
1.3. Немного о словах: каким бывает машинное обучение	17
1.4. Особенности человеческого мозга.	21
1.5. Пределы нейробиологии: что мы на самом деле знаем?	26
1.6. Блеск и нищета современных нейронных сетей	30
Глава 2. Предварительные сведения, или Курс молодого бойца	38
2.1. Теорема Байеса	39
2.2. Функции ошибки и регуляризация.	53
2.3. Расстояние Кульбака — Лейблера и перекрестная энтропия.	63
2.4. Градиентный спуск: основы	69
2.5. Граф вычислений и дифференцирование на нем	75
2.6. И о практике: введение в TensorFlow и Keras.	81
Глава 3. Перцептрон, или Эмбрион мудрого компьютера	93
3.1. Когда появились искусственные нейронные сети	94
3.2. Как работает перцептрон	97
3.3. Современные перцептроны: функции активации.	105
3.4. Как же обучаются настоящие нейроны	113
3.5. Глубокие сети: в чем прелесть и в чем сложность?	117
3.6. Пример: распознавание рукописных цифр на TensorFlow	123

Часть II. Основные архитектуры

Глава 4. Быстрее, глубже, сильнее, или Об оврагах, долинах и грамплинах	137
4.1. Регуляризация в нейронных сетях	138
4.2. Как инициализировать веса	142
4.3. Нормализация по мини-батчам.	153
4.4. Метод моментов: Ньютон, Нестеров и Гессе	164
4.5. Адаптивные варианты градиентного спуска	169
Глава 5. Сверточные нейронные сети и автокодировщики, или Не верь глазам своим	176
5.1. Зрительная кора головного мозга	177
5.2. Свертки и сверточные сети	182
5.3. Свертки для распознавания цифр	199
5.4. Современные сверточные архитектуры.	206
5.5. Автокодировщики	214
5.6. Пример: кодируем рукописные цифры	219

Глава 6. Рекуррентные нейронные сети, или Как правильно кусать себя за хвост231
6.1. Мотивация: обработка последовательностей232
6.2. Распространение ошибки и архитектуры RNN236
6.3. LSTM242
6.4. GRU и другие варианты249
6.5. SCRN и другие: долгая память в обычных RNN253
6.6. Пример: порождаем текст символ за символом259

Часть III. Новые архитектуры и применения

Глава 7. Как научить компьютер читать, или Математик – Мужчина + Женщина =278
7.1. Интеллектуальная обработка текстов279
7.2. Распределенные представления слов: word2vec285
7.3. Русскоязычный word2vec на практике297
7.4. GloVe: раскладываем матрицу правильно305
7.5. Вверх и вниз от представлений слов313
7.6. Рекурсивные нейронные сети и синтаксический разбор322

Глава 8. Современные архитектуры, или Как в споре рождается истина330
8.1. Модели с вниманием и encoder-decoder331
8.2. Порождающие модели и глубокое обучение341
8.3. Состязательные сети348
8.4. Практический пример и трюк с логистическим сигмоидом353
8.5. Архитектуры, основанные на GAN359

Глава 9. Глубокое обучение с подкреплением, или Удивительное происшествие с чемпионом372
9.1. Обучение с подкреплением373
9.2. Марковские процессы принятия решений379
9.3. От TDGammon к DQN391
9.4. Бамбуковая хлопущка399
9.5. Градиент по стратегиям и другие применения405

Глава 10. Нейробайесовские методы, или Прошлое и будущее машинного обучения409
10.1. Теорема Байеса и нейронные сети410
10.2. Алгоритм EM412
10.3. Вариационные приближения419
10.4. Вариационный автокодировщик426
10.5. Байесовские нейронные сети и дропаут438
10.6. Заключение: что не вошло в книгу и что будет дальше446

Благодарности450
-------------------------	------

Литература451
----------------------	------