

УДК 004.85Deeplearning4j  
ББК 32.813  
П20

**Паттерсон Дж., Гибсон А.**  
П20 Глубокое обучение с точки зрения практика / пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2018. – 418 с.: ил.

**ISBN 978-5-97060-481-6**

Интерес к машинному обучению зашкаливает, но завышенные ожидания нередко губят проекты еще на ранней стадии. Как машинное обучение — и особенно глубокие нейронные сети — может изменить вашу организацию? Эта книга не только содержит практически полезную информацию о предмете, но и поможет приступить к созданию эффективных сетей глубокого обучения.

Авторы сначала раскрывают фундаментальные вопросы глубокого обучения — настройка, распараллеливание, векторизация, конвейеры операций, а затем переходят к библиотеке Deeplearning4j (DL4J), предназначенной для разработки технологических процессов профессионального уровня. На реальных примерах читатель познакомится с методами и стратегиями обучения глубоких сетей с различной архитектурой и их распараллеливания в кластерах Hadoop и Spark.

Издание предназначено для специалистов по анализу данных, находящихся в поисках более широкого и практического понимания принципов глубокого обучения.

УДК 004.85Deeplearning4j  
ББК 32.813

Authorized Russian translation of the English edition of Deep Learning, ISBN 9781491914250.  
© 2017 Josh Patterson, Adam Gibson.

This translation is published and sold by permission of O'Reilly Media, Inc., which owns or controls all rights to publish and sell the same.

Все права защищены. Любая часть этой книги не может быть воспроизведена в какой бы то ни было форме и какими бы то ни было средствами без письменного разрешения владельцев авторских прав.

ISBN 978-1-491-91425-0 (англ.)  
ISBN 978-5-97060-481-6 (рус.)

Copyright © 2017 Josh Patterson, Adam Gibson  
© Оформление, издание, перевод, ДМК Пресс, 2018

# Содержание

<b>Предисловие</b> .....	14
<b>Глава 1. Обзор машинного обучения</b> .....	20
Обучающиеся машины.....	20
Как машины могут обучаться? .....	21
Биологические корни.....	23
Что такое глубокое обучение? .....	24
Вниз по кроличьей норе .....	25
Формулировка вопросов .....	26
Математические основания машинного обучения: линейная алгебра.....	26
Скаляры.....	26
Векторы .....	26
Матрицы .....	27
Тензоры.....	27
Гиперплоскости .....	28
Математические операции.....	28
Преобразование данных в векторы .....	28
Решение систем уравнений.....	30
Математические основания машинного обучения: статистика .....	32
Вероятность .....	32
Условные вероятности .....	34
Апостериорная вероятность.....	34
Распределения вероятности .....	35
Выборка и генеральная совокупность .....	37
Методы с перевыборкой .....	37
Смещение выборки .....	37
Правдоподобие .....	38
Как работает машинное обучение? .....	38
Регрессия.....	38
Классификация .....	40
Кластеризация .....	40
Недообучение и переобучение.....	41
Оптимизация.....	41
Выпуклая оптимизация .....	42
Градиентный спуск.....	43
Стохастический градиентный спуск.....	45
Квазиньютоновские методы оптимизации.....	46
Порождающие и дискриминантные модели.....	46
Логистическая регрессия .....	47
Логистическая функция .....	48
Интерпретация результата логистической регрессии.....	48
Оценивание модели .....	49
Матрица неточностей .....	49
Итоги .....	52

<b>Глава 2. Основы нейронных сетей и глубокого обучения</b>	<b>53</b>
Нейронные сети	53
Биологический нейрон	55
Перцептрон	57
Многослойные сети прямого распространения	60
Обучение нейронных сетей	64
Обучение методом обратного распространения	65
Функции активации	71
Линейная функция	71
Сигмоида	72
Функция tanh	73
Функция hardtanh	73
Функция softmax	73
Линейная ректификация	74
Функции потерь	76
Применяемые обозначения	76
Функции потерь для регрессии	77
Функции потерь для классификации	78
Функции потерь для реконструкции	80
Гиперпараметры	81
Скорость обучения	81
Регуляризация	82
Импульс	82
Разреженность	82
<b>Глава 3. Основания глубоких сетей</b>	<b>83</b>
Определение глубокого обучения	83
Что такое глубокое обучение?	83
Как организована эта глава	92
Общие архитектурные принципы глубоких сетей	92
Параметры	93
Слои	93
Функции активации	94
Функции потерь	95
Алгоритмы оптимизации	96
Гиперпараметры	98
Итоги	102
Строительные блоки глубоких сетей	102
Ограниченные машины Больцмана	103
Автокодировщики	108
Вариационные автокодировщики	109
<b>Глава 4. Основные архитектуры глубоких сетей</b>	<b>111</b>
Сети, предобученные без учителя	111
Глубокие сети доверия	111
Порождающие состязательные сети	114
Сверточные нейронные сети (СНС)	117
Биологические корни	118
Интуитивное описание	119
Общий взгляд на архитектуру СНС	120
Входной слой	121

Пулинговые слои .....	128
Полносвязные слои .....	129
Другие применения СНС .....	129
Самые известные СНС .....	130
Итоги .....	130
<b>Рекуррентные нейронные сети .....</b>	<b>131</b>
Моделирование времени .....	131
Трехмерный вход .....	133
Почему не марковские модели? .....	135
Общая архитектура рекуррентной нейронной сети .....	136
LSTM-сети .....	136
Предметно-ориентированные приложения и гибридные сети .....	143
<b>Рекурсивные нейронные сети .....</b>	<b>144</b>
Архитектура сети .....	144
Разновидности рекурсивных нейронных сетей .....	145
Применение рекурсивных нейронных сетей .....	145
<b>Итоги и обсуждение .....</b>	<b>145</b>
Приведет ли глубокое обучение к отмиранию всех прочих алгоритмов? .....	146
Оптимальный метод зависит от задачи .....	146
Когда мне может понадобиться глубокое обучение? .....	147
<b>Глава 5. Построение глубоких сетей .....</b>	<b>148</b>
<b>Выбор глубокой сети для решения задачи .....</b>	<b>148</b>
Табличные данные и многослойные перцептроны .....	148
Изображения и сверточные нейронные сети .....	149
Последовательности, временные ряды и рекуррентные нейронные сети .....	150
Применение гибридных сетей .....	151
<b>Инструментарий DL4J .....</b>	<b>151</b>
Векторизация и DataVec .....	152
Среды выполнения и ND4J .....	152
<b>Основные концепции DL4J API .....</b>	<b>153</b>
Загрузка и сохранение моделей .....	153
Получение входных данных для модели .....	154
Задание архитектуры модели .....	154
Обучение и оценивание .....	155
<b>Моделирование CSV-данных с помощью многослойных перцептронов .....</b>	<b>156</b>
Подготовка входных данных .....	158
Задание архитектуры сети .....	159
Обучение модели .....	161
Оценивание модели .....	161
<b>Моделирование рукописных цифр с помощью СНС .....</b>	<b>162</b>
Реализация СНС LeNet на Java .....	163
Загрузка и векторизация входных изображений .....	165
Архитектура сети LeNet в DL4J .....	165
Обучение СНС .....	168
<b>Моделирование последовательных данных с помощью рекуррентной нейронной сети .....</b>	<b>169</b>
Порождение текста в стиле Шекспира с помощью LSTM-сети .....	169
Классификация временных рядов, содержащих показания датчика, с помощью LSTM-сети .....	177
<b>Применение автокодировщиков для обнаружения аномалий .....</b>	<b>183</b>

Java-программа автокодировщика .....	183
Подготовка входных данных .....	187
Архитектура и обучение сети автокодировщика.....	187
Оценивание модели .....	188
Использование вариационных автокодировщиков для реконструкции цифр из набора MNIST .....	189
Программа реконструкции цифр для набора MNIST .....	190
Изучение модели VAE .....	192
Применение глубокого обучения в обработке естественного языка.....	195
Обучение погружениям слов с помощью Word2Vec .....	196
Распределенные представления предложений с помощью векторов абзацев.....	201
Применение векторов абзацев для классификации документов .....	204
<b>Глава 6. Настройка глубоких сетей .....</b>	<b>209</b>
Основные концепции настройки глубоких сетей .....	209
Интуитивное описание построения глубоких сетей .....	209
Преобразование интуитивных представлений в пошаговую процедуру .....	210
Подбор сетевой архитектуры, соответствующей входным данным .....	211
Итоги .....	212
Соотнесение назначения модели с выходным слоем .....	213
Выходной слой регрессионной модели .....	213
Выходной слой модели классификации .....	213
Количество слоев, количество параметров и объем памяти .....	216
Многослойные нейронные сети прямого распространения.....	216
Управление количеством слоев и параметров .....	217
Оценка требований к объему памяти.....	219
Стратегии инициализации весов .....	220
Ортогональная инициализация весов в РНС.....	221
Выбор функции активации.....	221
Сводная таблица функций активации .....	223
Применение функций потерь.....	223
Скорость обучения.....	225
Использование отношения обновлений к параметрам .....	226
Конкретные рекомендации по заданию скорости обучения .....	227
Как разреженность влияет на обучение.....	228
Применение методов оптимизации .....	229
Рекомендации по применению CGC .....	230
Применение распараллеливания и GPU для ускорения обучения.....	231
Онлайновое обучение и параллельные итеративные алгоритмы.....	232
Распараллеливание CGC в DL4J.....	234
Графические процессоры.....	236
Управление периодами и размером мини-пакета .....	237
Компромиссы при определении размера мини-пакета.....	238
О применении регуляризации .....	239
Априорная функция как регуляризатор .....	239
Регуляризация по максимальной норме .....	240
Прореживание .....	240
Другие вопросы, связанные с регуляризацией .....	242
Дисбаланс классов .....	242
Методы выборки из классов .....	244
Взвешенные функции потерь.....	244

Борьба с переобучением .....	245
Использование статистики сети из интерфейса настройки .....	246
Обнаружение неудачной инициализации весов.....	248
Обнаружение неперемешанных данных.....	249
Обнаружение проблем, относящихся к регуляризации .....	251
<b>Глава 7. Настройка глубоких сетей с конкретной архитектурой .....</b>	<b>253</b>
Сверточные нейронные сети (СНС) .....	253
Общие архитектурные паттерны сверточных сетей .....	254
Конфигурирование сверточных слоев .....	257
Конфигурирование пулинговых слоев .....	261
Перенос обучения.....	262
Рекуррентные нейронные сети .....	263
Входные данные и входной слой сети .....	264
Выходные слои и RnnOutputLayer .....	264
Обучение сети.....	265
Отладка типичных проблем в LSTM .....	267
Дополнение и маскирование.....	267
Применение маскирования для оценивания и скоринга.....	268
Варианты архитектуры рекуррентных сетей .....	269
Ограниченные машины Больцмана.....	269
Скрытые блоки и моделирование доступной информации .....	270
Типы блоков.....	271
Регуляризация в ОМБ.....	271
Глубокие сети доверия .....	272
Применение импульса .....	272
Применение регуляризации.....	273
Задание числа скрытых блоков .....	273
<b>Глава 8. Векторизация .....</b>	<b>274</b>
Введение в векторизацию в машинном обучении.....	274
Зачем нужно векторизовать данные?.....	275
Стратегии обработки табличных исходных данных.....	277
Конструирование признаков и методы нормировки .....	279
Применение библиотеки DataVec для ETL и векторизации .....	285
Векторизация изображений .....	286
Представление изображений в DL4J .....	286
Нормировка данных изображения с помощью DataVec.....	288
Векторизация последовательных данных .....	289
Основные виды источников последовательных данных .....	289
Векторизация последовательных данных с помощью DataVec .....	290
Векторизация текста .....	294
Мешок слов .....	295
TF-IDF .....	296
Сравнение Word2Vec и векторной модели .....	299
Работа с графами .....	300
<b>Глава 9. Глубокое обучение и DL4J на платформе Spark .....</b>	<b>301</b>
Введение в использование DL4J совместно с Spark и Hadoop .....	301
Запуск Spark из командной строки .....	303
Конфигурирование и настройка Spark.....	305

Выполнение Spark в кластере Mesos .....	306
Выполнение Spark поверх YARN.....	307
Общее руководство по настройке Spark .....	309
Настройка задач DL4J для Spark .....	311
Подготовка проекта Maven для Spark и DL4J .....	312
Шаблон секции зависимостей в файле pom.xml .....	314
Настройка POM-файла для CDH 5.X.....	317
Настройка POM-файла для HDP 2.4 .....	317
Отладка Spark и Hadoop .....	318
Типичные проблемы при работе с ND4J.....	318
Параллельное выполнение DL4J на платформе Spark .....	319
Минимальный пример обучения на платформе Spark.....	320
Рекомендации по использованию DL4J API для Spark.....	322
Пример многослойного перцептрона на платформе Spark .....	323
Конфигурирование архитектуры МСП для Spark.....	326
Распределенное обучение и оценивание модели .....	327
Создание и выполнение задачи Spark .....	328
Порождение текстов в стиле Шекспира с помощью Spark и LSTM-сети .....	328
Конфигурирование архитектуры LSTM-сети .....	330
Обучение, наблюдение за ходом работы и интерпретация результатов .....	331
Моделирование набора MNIST с помощью сверточной нейронной сети в кластере Spark .....	332
Конфигурирование задачи Spark и загрузка набора данных MNIST .....	334
Конфигурирование и обучение CHC LeNet.....	335
<b>Приложение А. Что такое искусственный интеллект? .....</b>	<b>337</b>
Положение на данный момент .....	338
Определение глубокого обучения .....	338
Определение искусственного интеллекта .....	338
Зима не за горами.....	345
<b>Приложение В. RL4J и обучение с подкреплением .....</b>	<b>347</b>
Введение.....	347
Марковский процесс принятия решений .....	347
Терминология .....	348
Различные варианты .....	349
Безмодельное обучение .....	349
Наблюдаемое состояние .....	349
Однопользовательские и состязательные игры .....	349
Q-обучение.....	350
От политики к нейронным сетям.....	350
Перебор политик .....	352
Исследование и использование.....	355
Уравнение Беллмана .....	356
Выборка начальных состояний .....	357
Реализация Q-обучения .....	358
Моделирование Q(s,a) .....	359
Воспроизведение опыта .....	359
Сверточные слои и предварительная обработка изображений.....	360
Обработка истории.....	361
Двойное Q-обучение .....	361

Ограничение .....	362
Масштабирование вознаграждений .....	362
Приоритетное воспроизведение .....	362
График, визуализация и среднее значение Q .....	362
RL4J .....	365
Заключение .....	366
<b>Приложение С. Числа, которые должен знать каждый .....</b>	<b>367</b>
<b>Приложение D. Нейронные сети и обратное распространение:</b>	
<b>математическое описание .....</b>	<b>368</b>
Введение .....	368
Обратное распространение в многослойном перцептроне .....	369
<b>Приложение E. ND4J API .....</b>	<b>372</b>
Дизайн и основы использования .....	372
Что такое NDArray? .....	373
Общий синтаксис ND4J .....	374
Основы работы с массивами NDArray .....	375
Класс Dataset .....	377
Создание входных векторов .....	378
Основы создания векторов .....	378
Класс MLLibUtil .....	379
Преобразование INDArray в вектор MLLib .....	379
Преобразование вектора MLLib в INDArray .....	379
Получение предсказаний от модели в DL4J .....	380
Совместное использование DL4J и ND4J .....	380
<b>Приложение F. Библиотека DataVec .....</b>	<b>382</b>
Загрузка данных для машинного обучения .....	382
Загрузка CSV-данных для многослойного перцептрона .....	384
Загрузка изображений для сверточной нейронной сети .....	385
Загрузка последовательных данных для рекуррентных нейронных сетей .....	386
Подготовка данных средствами DataVec .....	387
Преобразования DataVec: основные понятия .....	388
Преобразования DataVec: пример .....	389
<b>Приложение G. Работа с DL4J на уровне исходного кода .....</b>	<b>391</b>
Проверка, установлен ли Git .....	391
Клонирование основных проектов, связанных с DL4J .....	391
Скачивание исходного кода в виде zip-файла .....	392
Сборка библиотеки из исходного кода с помощью Maven .....	392
<b>Приложение H. Подготовка проектов на базе DL4J .....</b>	<b>393</b>
Создание нового проекта на базе DL4J .....	393
Java .....	393
Работа с Maven .....	394
Интегрированные среды разработки (IDE) .....	395
Настройка других POM-файлов Maven .....	396
ND4J и Maven .....	396



<b>Приложение I. Подготовка проектов на базе DL4J к работе с GPU .....</b>	<b>397</b>
Переключение библиотек в режим работы с GPU .....	397
Выбор GPU.....	397
Обучение на системе с несколькими GPU .....	398
CUDA на разных платформах.....	398
Мониторинг производительности GPU.....	399
<b>Приложение J. Отладка проблем с установкой DL4J .....</b>	<b>400</b>
Предыдущая установка .....	400
Ошибки нехватки памяти при сборке из исходного кода .....	400
Старые версии Maven .....	400
Maven и переменная среды PATH.....	400
Недопустимые версии JDK.....	401
C++ и другие средства разработки.....	401
Windows и путь к включаемым файлам.....	401
Мониторинг GPU.....	401
Использование JVisualVM .....	401
Работа с Clojure .....	402
Поддержка чисел с плавающей точкой в OS X.....	402
Ошибка разветвления-соединения в Java 7 .....	402
Предостережения.....	402
Клонирование других репозиториев .....	402
Проверьте зависимости Maven.....	403
Переустановка зависимостей .....	403
Если ничего не помогает .....	403
Различные платформы.....	403
OS X .....	403
Windows.....	403
Linux .....	404
<b>Предметный указатель .....</b>	<b>405</b>
<b>Об авторах .....</b>	<b>416</b>
<b>Об иллюстрации на обложке .....</b>	<b>417</b>