

УДК 004.4
ББК 32.972
Х20

Монарх (Манро) Р.

Х20 Машинное обучение с участием человека / пер. с англ. В. И. Бахура. – М.: ДМК Пресс, 2022. – 498 с.: ил.

ISBN 978-5-97060-934-7

Эта книга нацелена на изучение взаимодействия искусственного интеллекта и человека в процессе создания и эксплуатации систем машинного обучения. В отличие от большинства курсов по машинному обучению, сосредоточенных на алгоритмах, большое внимание уделяется работе с данными: их маркировке, аннотированию, проверке и обновлению. Впервые под одной обложкой собраны наиболее распространенные стратегии аннотирования, активного обучения и смежных задач, таких как проектирование интерфейса для аннотирования.

Книга предназначена для специалистов по работе с данными, разработчиков программного обеспечения и тех, кто делает первые шаги в работе с машинным обучением.

УДК 004.4
ББК 32.972

Original English language edition published by Manning Publications USA. Copyright © 2021 by Manning Publications. Russian-language edition copyright © 2022 by DMK Press. All rights reserved.

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 978-1-6172-9674-1 (англ.)
ISBN 978-5-97060-934-7 (рус.)

© Manning Publications, 2021
© Перевод, оформление, издание, ДМК Пресс, 2022

Содержание

Предисловие	16
Введение	18
Благодарности	19
Об этой книге	21
Об авторе	25

Часть I ПЕРВЫЕ ШАГИ

1 Введение в машинное обучение с участием человека	27
1.1 Базовые принципы машинного обучения с участием человека	28
1.2 Введение в аннотирование	30
1.2.1 Простые и более сложные стратегии аннотирования	30
1.2.2 Устранение пробелов в области научных знаний о данных	30
1.2.3 Качество аннотирования человеком: почему это трудно?	31
1.3 Введение в активное обучение: повышение скорости и снижение стоимости обучающих данных	33
1.3.1 Три широкие стратегии отбора активного обучения: неопределенность, разнообразие и случайность	33
1.3.2 Что такое случайный выбор оценочных данных?	37
1.3.3 Когда использовать активное обучение?	38
1.4 Машинное обучение и взаимодействие человек–компьютер	40
1.4.1 Пользовательские интерфейсы: как вы создаете обучающие данные?	40
1.4.2 Прайминг: что может повлиять на человеческое восприятие?	42
1.4.3 Плюсы и минусы создания меток путем оценки прогнозов машинного обучения	43
1.4.4 Основные принципы проектирования интерфейсов аннотации	43
1.5 Машинное обучение в помощь человеку или машинное обучение с участием человека	43
1.6 Перенос обучения для запуска ваших моделей	44
1.6.1 Перенос обучения в компьютерном зрении	46
1.6.2 Перенос обучения при обработке естественного языка	46
1.7 Чего ожидать от этого текста	49
Резюме	50

2	Начало работы с машинным обучением с участием человека (human-in-the-loop)	52
2.1	За пределами активного обучения: ваш первый алгоритм активного обучения	53
2.2	Архитектура вашей первой системы	55
2.3	Интерпретация прогнозов модели и данных для активного обучения.....	59
2.3.1	Ранжирование достоверности	60
2.3.2	Выявление выбросов	61
2.3.3	Чего можно ожидать в процессе итераций.....	64
2.4	Построение интерфейса для сбора меток человека	66
2.4.1	Простой интерфейс для маркировки текста.....	66
2.4.2	Управление данными машинного обучения	69
2.5	Развертывание вашей первой системы машинного обучения с участием человека	69
2.5.1	Всегда в первую очередь собирайте данные для оценки	72
2.5.2	Каждая точка данных получает шанс	75
2.5.3	Выбор правильных стратегий для ваших данных.....	76
2.5.4	Переобучение модели и итерации.....	79
	Резюме.....	80

Часть II АКТИВНОЕ ОБУЧЕНИЕ..... 82

3	Выборка неопределенности	84
3.1	Интерпретация неопределенности в модели машинного обучения....	85
3.1.1	Для чего искать неопределенность в вашей модели?	86
3.1.2	Softmax и распределения вероятностей	88
3.1.3	Интерпретация успешности активного обучения	90
3.2	Алгоритмы для выборки неопределенности.....	90
3.2.1	Выборка с наименьшим доверием	92
3.2.2	Выборка по пределу уверенности	94
3.2.3	Соотношение выборок	95
3.2.4	Энтропия (энтропия классификации)	97
3.2.5	Глубокое погружение в энтропию	100
3.3	Определение случаев запутанности различных типов моделей	101
3.3.1	Выборка неопределенности с помощью логистической регрессии и моделей MaxEnt	101
3.3.2	Выборка неопределенности с помощью метода опорных векторов (SVM)	103
3.3.3	Выборка неопределенности с помощью байесовских моделей	104
3.3.4	Выборка неопределенности с помощью деревьев решений и случайных лесов.....	105
3.4	Измерение неопределенности по нескольким прогнозам	106
3.4.1	Выборка неопределенности с помощью ансамбля моделей.....	106
3.4.2	Запрос по комитету и отсеивание.....	108
3.4.3	Разница между алеаторной и эпистемической неопределенностями	110
3.4.4	Классификация с несколькими метками и непрерывными значениями	111
3.5	Определение правильного числа элементов для проверки человеком	112

3.5.1	Выборка неопределенности с ограниченным бюджетом	113
3.5.2	Выборка неопределенности с временными ограничениями	114
3.5.3	Когда остановиться, если нет ограничений по времени или бюджету?	115
3.6	Оценка успешности активного обучения	115
3.6.1	Нужны ли мне новые тестовые данные?	115
3.6.2	Нужны ли мне новые данные для проверки?	116
3.7	Памятка по выборке неопределенности	118
3.8	Дополнительная литература	120
3.8.1	Дополнительная литература по наименее достоверной выборке	121
3.8.2	Дополнительная литература по выборке с пределом достоверности	121
3.8.3	Дополнительная литература по доверительной выборке	121
3.8.4	Дополнительная литература по выборке на основе энтропии	121
3.8.5	Дополнительная литература по другим моделям машинного обучения	122
3.8.6	Дополнительная литература по выборке неопределенности на основе ансамблей	122
	Резюме	123

4 Выборка разнообразия

4.1	Осознание того, чего вы не знаете: выявление пробелов в знаниях вашей модели	126
4.1.1	Пример данных для выборки разнообразия	129
4.1.2	Интерпретация нейронных моделей для выборки разнообразия	130
4.1.3	Получение информации из скрытых слоев в PyTorch	132
4.2	Выборка выбросов на основе модели	135
4.2.1	Использование данных проверки для ранжирования активаций	136
4.2.2	Какие слои следует использовать для расчета выбросов модели?	140
4.2.3	Ограничения выбросов на данных моделей	141
4.3	Кластерная выборка	142
4.3.1	Состав кластера, центроиды и выбросы	143
4.3.2	Любой из существующих во вселенной алгоритмов кластеризации	144
4.3.3	Кластеризация k-средних с косинусным сходством	146
4.3.4	Уменьшение размерности параметров с помощью вложений или анализа главных компонент	149
4.3.5	Другие алгоритмы кластеризации	151
4.4	Репрезентативная выборка	153
4.4.1	Репрезентативная выборка нечасто используется обособленно	154
4.4.2	Простая репрезентативная выборка	156
4.4.3	Адаптивная репрезентативная выборка	157
4.5	Выборка для получения реального разнообразия	159
4.5.1	Распространенные проблемы разнообразия обучающих данных	160
4.5.2	Стратифицированная выборка для обеспечения разнообразия демографических данных	162
4.5.3	Представленный и представляющий: что важно?	163
4.5.4	Демографическая точность	164
4.5.5	Ограничения выборки для определения реального разнообразия	165
4.6	Выборка разнообразия с различными типами моделей	166

4.6.1	Выбросы на основе различных типов моделей.....	166
4.6.2	Кластеризация с использованием различных типов моделей	166
4.6.3	Репрезентативная выборка с различными типами моделей	167
4.6.4	Выборка для реального разнообразия с различными типами моделей	167
4.7	Краткая памятка по выборке разнообразия.....	167
4.8	Дополнительная литература.....	169
4.8.1	Дополнительная литература по выбросам на основе моделей	169
4.8.2	Дополнительная литература по кластерной выборке	169
4.8.3	Дополнительная литература по репрезентативной выборке.....	170
4.8.4	Дополнительная литература по выборке для реального разнообразия	170
Резюме	171

5	Расширенное активное обучение	173
5.1	Сочетание выборки неопределенности и выборки разнообразия	173
5.1.1	Выборка наименьшего доверия с выборкой на основе кластеров ...	174
5.1.2	Выборка неопределенности с выбросами по модели.....	177
5.1.3	Выборка неопределенности с выбросами по модели и кластеризацией	179
5.1.4	Репрезентативная выборка на основе кластерной выборки.....	179
5.1.5	Выборка из кластера с наибольшей энтропией	182
5.1.6	Другие комбинации стратегий активного обучения.....	185
5.1.7	Сочетание результатов активного обучения	186
5.1.8	Выборка для уменьшения предполагаемой ошибки	187
5.2	Активный перенос обучения для выборки неопределенности	189
5.2.1	Учим модель предсказывать собственные ошибки	190
5.2.2	Применение активного переноса обучения	191
5.2.3	Активный перенос обучения с большим количеством слоев	194
5.2.4	Плюсы и минусы активного переноса обучения.....	195
5.3	Применение активного переноса обучения к репрезентативной выборке	196
5.3.1	Использование модели для предсказания неизвестного.....	196
5.3.2	Активный перенос обучения для адаптивной репрезентативной выборки.....	198
5.3.3	Плюсы и минусы активного переноса обучения для репрезентативной выборки.....	199
5.4	Активный перенос обучения для адаптивной выборки.....	200
5.4.1	Адаптация выборки неопределенности посредством прогнозирования неопределенности.....	200
5.4.2	Плюсы и минусы метода ATLAS	203
5.5	Краткие памятки по расширенному активному обучению	204
5.6	Дополнительная литература по активному переносу обучения	206
Резюме	207

6	Активное обучение для решения различных задач машинного обучения	208
6.1	Использование активного обучения для обнаружения объектов.....	209
6.1.1	Точность выявления объектов: достоверность меток и локализация	211

6.1.2	Выборка неопределенности для оценки достоверности меток и локализации при выявлении объектов	213
6.1.3	Выборка разнообразия для достоверности меток и локализации при выявлении объектов	215
6.1.4	Активный перенос обучения для распознавания объектов	219
6.1.5	Низкий порог распознавания объектов во избежание закрепления необъективности	219
6.1.6	Создание образцов обучающих данных для репрезентативной выборки, схожих с прогнозами	221
6.1.7	Выборка разнообразия по изображениям при распознавании объектов	222
6.1.8	Создание более точных масок при использовании многоугольников	223
6.2	Использование активного обучения для семантической сегментации	224
6.2.1	Точность семантической сегментации	225
6.2.2	Выборка неопределенности для семантической сегментации	227
6.2.3	Выборка разнообразия для семантической сегментации	228
6.2.4	Активный перенос обучения для семантической сегментации	229
6.2.5	Выборка разнообразия по изображениям для семантической сегментации	229
6.3	Применение активного обучения для маркировки последовательностей	230
6.3.1	Точность маркировки последовательностей	231
6.3.2	Выборка неопределенности для маркировки последовательностей	232
6.3.3	Выборка разнообразия для маркировки последовательностей	233
6.3.4	Активный перенос обучения для маркировки последовательностей	236
6.3.5	Стратифицированная выборка по достоверности и токенам	237
6.3.6	Создание образцов обучающих данных для репрезентативной выборки, похожих на ваши прогнозы	237
6.3.7	Маркировка всей последовательности	237
6.3.8	Выборка разнообразия по документу при маркировке последовательностей	238
6.4	Применение активного обучения для генерации языка	238
6.4.1	Вычисление точности для систем генерации языка	239
6.4.2	Выборка неопределенности для генерации языка	240
6.4.3	Выборка разнообразия для генерации языка	241
6.4.4	Активный перенос обучения для генерации языка	242
6.5	Применение активного обучения к другим задачам машинного обучения	242
6.5.1	Активное обучение для поиска информации	243
6.5.2	Активное обучение для видео	245
6.5.3	Активное обучение для речи	246
6.6	Выбор подходящего количества элементов для проверки человеком	247
6.6.1	Активная разметка полностью или частично аннотированных данных	247
6.6.2	Совмещение машинного обучения с аннотированием	248
6.7	Дополнительная литература	248
	Резюме	249

Часть III АННОТИРОВАНИЕ.....250

7	Работа с людьми, аннотирующими ваши данные	252
7.1	Введение в аннотирование.....	254
7.1.1	Три правила хорошего аннотирования данных.....	255
7.1.2	Аннотирование данных и проверка прогнозов модели.....	256
7.1.3	Аннотации человека, полученные в процессе машинного обучения.....	256
7.2	Штатные эксперты.....	257
7.2.1	Зарботная плата для штатных сотрудников.....	258
7.2.2	Защищенность штатных сотрудников.....	259
7.2.3	Вовлеченность штатных сотрудников.....	259
7.2.4	Совет: всегда проводите сеансы аннотирования своими силами.....	261
7.3	Сотрудники на аутсорсинге.....	263
7.3.1	Зарплата для аутсорсинговых работников.....	264
7.3.2	Защищенность аутсорсинговых работников.....	266
7.3.3	Вовлеченность аутсорсинговых работников.....	266
7.3.4	Совет: общайтесь с вашими аутсорсинговыми сотрудниками.....	267
7.4	Краудсорсинговые работники.....	268
7.4.1	Зарплата для сотрудников краудсорсинга.....	270
7.4.2	Защищенность краудсорсинговых работников.....	271
7.4.3	Вовлеченность краудсорсинговых работников.....	272
7.4.4	Совет: создайте условия для стабильной работы и карьерного роста.....	273
7.5	Другие виды рабочей силы.....	273
7.5.1	Конечные пользователи.....	274
7.5.2	Волонтеры.....	275
7.5.3	Любители игр.....	277
7.5.4	Прогноз модели в качестве аннотации.....	278
7.6	Оценка требуемого объема аннотирования.....	280
7.6.1	Уравнение порядка количества необходимых аннотаций.....	280
7.6.2	От одной до четырех недель на обучение аннотированию и уточнение заданий.....	282
7.6.3	Для оценки затрат используйте пилотные аннотации и показатели точности.....	283
7.6.4	Сочетание разных типов трудовых ресурсов.....	283
	Резюме.....	284
8	Контроль качества при аннотировании данных	285
8.1	Сравнение аннотаций с истинными значениями ответов.....	286
8.1.1	Согласие аннотатора с базовыми истинными данными.....	289
8.1.2	Какой базовый уровень использовать для ожидаемой точности?.....	292
8.2	Межаннотаторское согласие.....	293
8.2.1	Введение в межаннотаторское согласие.....	294
8.2.2	Преимущества вычисления межаннотаторского согласия.....	296
8.2.3	Согласие по набору данных с помощью альфы Криппендорфа.....	299
8.2.4	Для чего, помимо маркировки, применима альфа Криппендорфа.....	303
8.2.5	Индивидуальное согласие аннотаторов.....	304

8.2.6	Согласие по каждой метке и каждому демографическому показателю	308
8.2.7	Повышение точности с помощью согласия для реального разнообразия	309
8.3	Агрегирование аннотаций для создания обучающих данных	309
8.3.1	Агрегирование аннотаций при общем согласии	310
8.3.2	Математический расчет для несогласных аннотаторов и низкого уровня согласия	311
8.3.3	Агрегирование аннотаций при несогласии аннотаторов	312
8.3.4	Достоверность с подачи аннотатора	314
8.3.5	Решаем, каким меткам доверять: неопределенность аннотации	315
8.4	Контроль качества посредством экспертной оценки	318
8.4.1	Набор и обучение квалифицированных сотрудников	319
8.4.2	Обучение персонала до уровня экспертов	320
8.4.3	Экспертиза с помощью машинного обучения	320
8.5	Многоэтапные рабочие процессы и задачи рецензирования	321
8.6	Дополнительная литература	323
	Резюме	324

9 Углубленное аннотирование и дополнение данных

9.1	Качественное аннотирование для субъективных задач	326
9.1.1	Выяснение предположений аннотаторов	329
9.1.2	Определение приемлемых меток для субъективных задач	330
9.1.3	Доверие к аннотатору для анализа разнообразия ответов	332
9.1.4	Байесовская сыворотка правды для субъективных суждений	334
9.1.5	Встраивание простых задач в более сложные	336
9.2	Машинное обучение для контроля качества аннотаций	337
9.2.1	Расчет достоверности аннотации как задачи оптимизации	338
9.2.2	Согласование достоверности меток при разногласиях аннотаторов	339
9.2.3	Прогнозирование достоверности отдельной аннотации	342
9.2.4	Прогнозирование согласованности для отдельной аннотации	344
9.2.5	Определение аннотатора как бота	344
9.3	Предсказание модели в качестве аннотаций	345
9.3.1	Доверие к аннотациям на основе достоверных предсказаний модели	346
9.3.2	Использование прогнозов модели в качестве единого аннотатора	349
9.3.3	Перекрестная валидация для поиска ошибочно маркированных данных	350
9.4	Вложения и контекстуальные отображения	350
9.4.1	Обучение переноса из существующей модели	353
9.4.2	Представления из смежных легко аннотируемых задач	354
9.4.3	Метод самоконтроля: использование меток, присущих данным	355
9.5	Системы на основе поиска и системы на основе правил	357
9.5.1	Фильтрация данных с помощью правил	358
9.5.2	Поиск обучающих данных	359
9.5.3	Маскированная фильтрация характеристик	359
9.6	Легкий надзор над неконтролируемыми моделями	360

9.6.1	Адаптация неконтролируемой модели к контролируемой модели	360
9.6.2	Исследовательский анализ данных под контролем человека	362
9.7	Синтетические данные, создание данных и их дополнение	362
9.7.1	Синтетические данные	362
9.7.2	Создание данных	363
9.7.3	Дополнение данных	365
9.8	Внедрение информации об аннотациях в модели машинного обучения	365
9.8.1	Фильтрация, или взвешивание элементов по доверию к их меткам	366
9.8.2	Включение идентификации аннотатора во входные данные	366
9.8.3	Внедрение неопределенности в функцию потерь	367
9.9	Дополнительная литература по расширенному аннотированию	368
9.9.1	Дополнительная литература по субъективным данным	368
9.9.2	Дополнительная литература по машинному обучению для контроля качества аннотаций	368
9.9.3	Дополнительная литература по вложениям / контекстным представлениям	369
9.9.4	Дополнительная литература по системам на основе правил	370
9.9.5	Дополнительная литература по включению неопределенности аннотаций в последующие модели	370
	Резюме	371

10 Качественные аннотации для различных задач машинного обучения

10.1	Качество аннотаций для непрерывных задач	374
10.1.1	Базовая истина для непрерывных задач	374
10.1.2	Соглашение для непрерывных задач	375
10.1.3	Субъективность в непрерывных задачах	376
10.1.4	Агрегирование непрерывных оценок для создания обучающих данных	377
10.1.5	Машинное обучение для агрегирования непрерывных задач с целью создания обучающих данных	379
10.2	Качество аннотаций для задач распознавания объектов	381
10.2.1	Базовая истина для распознавания объектов	382
10.2.2	Согласие при распознавании объектов	384
10.2.3	Размерность и точность при распознавании объектов	385
10.2.4	Субъективность при распознавании объектов	386
10.2.5	Агрегирование аннотаций объектов для создания обучающих данных	386
10.2.6	Машинное обучение для аннотаций объектов	388
10.3	Качество аннотаций для семантической сегментации	389
10.3.1	Базовая истина для аннотации семантической сегментации	390
10.3.2	Соглашение для семантической сегментации	391
10.3.3	Субъективность аннотаций семантической сегментации	391
10.3.4	Агрегирование семантической сегментации для создания обучающих данных	392
10.3.5	Машинное обучение для агрегирования задач семантической сегментации при создании обучающих данных	393
10.4	Качество аннотации для маркировки последовательности	394

10.4.1	Базовая истина для маркировки последовательности	396
10.4.2	Базовая истина для маркировки последовательностей в реально непрерывных данных.....	397
10.4.3	Согласие по маркировке последовательностей.....	398
10.4.4	Машинное обучение и перенос обучения для маркировки последовательностей	398
10.4.5	Данные на основе правил, поиска и синтетических данных для маркировки последовательностей	401
10.5	Качество аннотаций для генерирования языковых материалов.....	401
10.5.1	Базовая истина для генерации языка	402
10.5.2	Согласие и агрегирование для генерации языка	403
10.5.3	Машинное обучение и обучение переноса для генерации языка	403
10.5.4	Синтетические данные для генерации языка.....	404
10.6	Качественное аннотирование для других задач машинного обучения.....	405
10.6.1	Аннотирование для поиска информации	405
10.6.2	Аннотирование для многоплановых задач.....	408
10.6.3	Аннотирование для видео	409
10.6.4	Аннотирование аудиоданных	410
10.7	Дополнительная литература по качеству аннотирования для различных задач машинного обучения	411
10.7.1	Дополнительная литература по компьютерному зрению.....	411
10.7.2	Дополнительная литература по аннотированию для обработки естественного языка	412
10.7.3	Дополнительная литература по аннотированию для информационного поиска.....	413
	Резюме.....	413

Часть IV ВЗАИМОДЕЙСТВИЕ ЧЕЛОВЕКА И КОМПЬЮТЕРА ПРИ МАШИННОМ ОБУЧЕНИИ.....415

11	Интерфейсы для аннотирования данных.....	417
11.1	Основные принципы взаимодействия человека и компьютера	418
11.1.1	Знакомство с доступностью, обратной связью и самостоятельностью	418
11.1.2	Проектирование интерфейсов для аннотирования.....	420
11.1.3	Сведение к минимуму движения глаз и прокрутки	421
11.1.4	Клавиатурные сочетания и устройства ввода	424
11.2	Эффективное нарушение правил.....	426
11.2.1	Прокрутка для пакетного аннотирования.....	426
11.2.2	Ножные педали	427
11.2.3	Голосовой ввод	427
11.3	Прайминг в интерфейсах аннотирования.....	428
11.3.1	Прайминг повторов.....	428
11.3.2	Где прайминг вреден.....	429
11.3.3	Где прайминг полезен	430
11.4	Сочетание интеллекта человека и машины	430
11.4.1	Обратная связь с аннотатором	431
11.4.2	Максимальная объективность за счет стороннего мнения.....	432

11.4.3	Преобразование непрерывных проблем в проблемы ранжирования	433
11.5	Интеллектуальные интерфейсы для максимальной отдачи человеческого интеллекта	435
11.5.1	Интеллектуальные интерфейсы для семантической сегментации	437
11.5.2	Интеллектуальные интерфейсы для распознавания объектов	440
11.5.3	Интеллектуальные интерфейсы для генерации языка	442
11.5.4	Интеллектуальные интерфейсы для маркировки последовательностей	445
11.6	Машинное обучение для содействия работе человека	447
11.6.1	Восприятие повышения эффективности	447
11.6.2	Активное обучение для повышения эффективности	448
11.6.3	Ошибки лучше их отсутствия для максимальной завершенности	449
11.6.4	Держите интерфейсы аннотирования отдельно от повседневных рабочих интерфейсов	450
11.7	Дополнительная литература	451
	Резюме	451

12 Продукты машинного обучения с участием человека

12.1	Определение продуктов для приложений машинного обучения с участием человека	454
12.1.1	Начните с решаемой вами задачи	454
12.1.2	Проектирование систем для решения задачи	455
12.1.3	Соединение Python и HTML	457
12.2	Пример 1: исследовательский анализ данных по заголовкам новостей	458
12.2.1	Предпосылки	459
12.2.2	Разработка и воплощение	460
12.2.3	Потенциальные расширения	461
12.3	Пример 2: сбор данных о событиях в области безопасности пищевых продуктов	462
12.3.1	Предпосылки	463
12.3.2	Разработка и реализация	464
12.3.3	Потенциальные расширения	465
12.4	Пример 3: идентификация велосипедов на изображениях	466
12.4.1	Предпосылки	466
12.4.2	Разработка и реализация	467
12.4.3	Потенциальные расширения	468
12.5	Дополнительная литература по созданию продуктов машинного обучения с участием человека	469
	Резюме	469
	Приложение. Краткое пособие по машинному обучению	470
	Предметный указатель	488