

УДК 004.89
ББК 32.813
Х98

Хултен Дж.

Х98 Разработка интеллектуальных систем / пер. с англ. В. С. Яценкова. – М.: ДМК Пресс, 2019. – 284 с.: ил.

ISBN 978-5-97060-760-2

Эта книга научит вас, как создавать интеллектуальные системы от начала до конца и использовать машинное обучение на практике. Вы узнаете, как эффективно применять свои навыки разработки программного обеспечения, науки о данных, машинного обучения и управления проектами.

Книга основана на более чем десятилетнем опыте создания интеллектуальных систем, которые обеспечивают сотни миллионов взаимодействий пользователей в день в некоторых из крупнейших и наиболее важных программных систем в мире.

Издание будет полезно инженерам-программистам, специалистам по машинному обучению и руководителям проектов, которые хотят создавать и внедрять эффективные интеллектуальные системы.

УДК 004.89
ББК 32.813

Authorized Russian translation of the English edition of Building Intelligent Systems: A Guide to Machine Learning Engineering ISBN 9781484234310 © 2018 by Geoff Hulten.

This translation is published and sold by permission of Apress Media, 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 978-1-4842-3431-0 (англ.)
ISBN 978-5-97060-760-2 (рус.)

© 2018 by Geoff Hulten
© Оформление, издание, перевод, ДМК Пресс, 2019

Содержание

Автор	15
Технический рецензент	16
Благодарности	17
Вступление	18
Часть I. ОСНОВЫ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ	23
Глава 1. Знакомство с интеллектуальными системами	24
Элементы интеллектуальной системы	25
Пример интеллектуальной системы	26
Умный тостер	26
Использование данных	27
Датчики и эвристический интеллект	29
Тостер с машинным обучением	30
Создание интеллектуальной системы	31
Итог главы	32
Темы для размышлений	32
Глава 2. Анализ применимости интеллектуальных систем	34
Типы задач, для которых нужны интеллектуальные системы	34
Масштабные задачи	35
Открытые задачи	35
Меняющиеся задачи	36
Сложные задачи	36
Ситуации, когда интеллектуальные системы успешны	37
Когда неполная система жизнеспособна и полезна	37
Обратная связь для улучшения интеллекта	38
Когда система способна влиять на результат	38
Когда интеллектуальная система окупается	39
Действительно ли нужна интеллектуальная система?	40
Итог главы	41
Темы для размышлений	41
Глава 3. Краткие основы работы с данными	42
Структурированные данные	42

Задавайте данным простые вопросы	44
Работа с моделями данных	45
Концепция машинного обучения	46
Распространенные ошибки при работе с данными	48
Нарушение доверительных интервалов	48
Зашумленность данных	48
Смещение данных	48
Устаревание данных	49
Необоснованное использование данных	49
Итог главы	49
Темы для размышлений	50

Глава 4. Определение целей интеллектуальной системы

Признаки хорошей цели	51
Пример затруднений при выборе цели	52
Типы целей	53
Организационные цели	54
Опережающие показатели	54
Результаты пользователя	56
Свойства модели	57
Расслоение целей	58
Способы измерить успех	58
Ожидание дополнительной информации	58
А/В-тестирование	59
Ручная маркировка	59
Опрос пользователей	60
Разделение задач	61
Сохраняйте актуальность целей	61
Итог главы	62
Темы для размышлений	62

Часть II. ИНТЕЛЛЕКТУАЛЬНЫЙ ОПЫТ

Глава 5. Компоненты интеллектуального опыта

Представление интеллекта пользователю	64
Пример представления интеллекта	66
Достижение целей системы	67
Пример достижения целей	68
Минимизация последствий ошибок интеллекта	69
Получение данных для расширения системы	70
Пример сбора данных	70
Итог главы	71
Темы для размышлений	72

Глава 6. Затруднения при разработке интеллектуального опыта	73
Интеллект делает ошибки	73
Безумные ошибки интеллекта	75
Интеллект совершает разные ошибки	76
Переменчивый интеллект	78
Человеческий фактор	79
Итог главы	80
Темы для размышлений	81
Глава 7. Разработка эффективного интеллектуального опыта	82
Действенность опыта	83
Частота взаимодействия	84
Выгода от взаимодействия	86
Цена взаимодействия	87
Обнаружение ошибки	87
Исправление ошибки	88
Качество интеллекта	89
Итог главы	90
Темы для размышлений	91
Глава 8. Режимы интеллектуального взаимодействия	92
Автоматизация действий	92
Запросы и подсказки	93
Организованная информация	94
Аннотации	96
Гибридный опыт	97
Итог главы	98
Темы для размышлений	99
Глава 9. Извлечение данных из опыта	100
Пример: TeamMaker	101
Прямое вмешательство	101
Увлекательное взаимодействие	102
Связь с результатами	102
Свойства хороших данных	103
Контекст, действия и результаты	103
Достоверный охват	104
Реальное применение	105
Отсутствие смещения	105
Отсутствие петель обратной связи	106
Масштаб	106
Правильное толкование данных	107
Скрытые наблюдения	108

Пользовательские рейтинги	108
Отчеты о проблемах	109
Эскалации	109
Пользовательские решения	109
Итог главы	110
Темы для размышлений	111
Глава 10. Проверка интеллектуального опыта	112
Достижение ожидаемого опыта	112
Работа с контекстом	113
Работа с интеллектом	114
Промежуточный итог	115
Достижение целей	115
Непрерывная проверка	116
Итоги главы	117
Темы для размышлений	118
Часть III. РЕАЛИЗАЦИЯ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ	119
Глава 11. Компоненты реализации интеллекта	120
Пример реализации интеллектуальной системы	120
Компоненты реализации интеллектуальной системы	123
Среда выполнения интеллекта	123
Распределение и доставка интеллекта	124
Канал интеллектуальной телеметрии	124
Среда разработки интеллекта	125
Оркестровка интеллекта	125
Итог главы	126
Темы для размышлений	126
Глава 12. Среда выполнения интеллекта	128
Сбор контекста	129
Извлечение признаков	130
Обновление моделей	131
Выполнение моделей	132
Результаты	132
Нестабильность в интеллекте	133
API интеллекта	133
Итог главы	134
Темы для размышлений	135
Глава 13. Где расположить интеллект?	136
Соображения по размещению интеллекта	136
Задержка при обновлении	137

Задержка выполнения.....	139
Стоимость эксплуатации	140
Автономная работа.....	142
Подходы к размещению интеллекта.....	142
Статический интеллект в составе продукта	142
Интеллект на стороне клиента	143
Интеллект на стороне сервера	144
Внутренний (кешируемый) интеллект.....	145
Гибридный интеллект	146
Итог главы	146
Темы для размышлений.....	147
Глава 14. Управление интеллектом	148
Механизм управления интеллектом.....	148
Сложная архитектура систем.....	149
Высокая частота обновления	149
Человеческий фактор	149
Проверка работоспособности интеллекта	150
Проверка на совместимость	150
Проверка ограничений на выполнение	151
Проверка на очевидные ошибки	151
Пробный запуск интеллекта	152
Однократное развертывание интеллекта.....	152
Тихий интеллект.....	153
Ограниченное развертывание.....	154
Флайтинг	154
Отмена обновления.....	155
Итог главы	156
Темы для размышлений.....	156
Глава 15. Интеллектуальная телеметрия	158
Зачем нужна телеметрия	158
Проверка текущей работоспособности.....	158
Проверка результатов пользователей.....	159
Сбор данных для развития интеллекта.....	160
Свойства эффективной телеметрии.....	161
Выборочное наблюдение	161
Резюмирование	162
Гибкий таргетинг	163
Общие проблемы	163
Смещение данных	163
Пропуск редких событий	164
Завышение значимости	165

Нарушение конфиденциальности.....	165
Итог главы	166
Темы для размышлений.....	167
Часть IV. СОЗДАНИЕ ИНТЕЛЛЕКТА.....	168
Глава 16. Общее представление об интеллекте	169
Пример интеллекта	169
Контексты.....	170
Реализация в среде выполнения	171
Доступность контекста для разработчика	172
Что может предсказать интеллект.....	173
Классификация	173
Оценка вероятности	174
Регрессия	175
Ранжирование.....	176
Составное предсказание	176
Итог главы	176
Темы для размышлений.....	177
Глава 17. Представление интеллекта	178
Критерии выбора представления интеллекта.....	178
Представление интеллекта в виде программы	179
Представление интеллекта в таблицах соответствий.....	180
Представление интеллекта в моделях	181
Линейные модели.....	182
Деревья решений.....	183
Нейронные сети.....	184
Итог главы	186
Темы для размышлений.....	186
Глава 18. Процесс создания интеллекта	187
Пример создания интеллекта	188
Понимание задачи и окружения	188
Определение критериев успеха.....	190
Получение данных.....	190
Данные для начального запуска.....	191
Данные из взаимодействий.....	192
Подготовка инструментов оценки	192
Простой эвристический интеллект	193
Машинное обучение.....	194
Поиск компромиссов.....	195
Оценка и повторение	195

Уровни зрелости интеллекта	196
Мастерство создания интеллекта	196
Анализ и отладка данных	197
Отладка на основе проверок и оценок	197
Интуитивное знание инструментария	197
Математика – нужна ли она?	198
Итог главы	198
Темы для размышлений	199
Глава 19. Оценка интеллекта	200
Оценка точности	200
Обобщение	201
Типы ошибок	201
Распределение ошибок	204
Оценка других типов прогнозов	204
Оценка регрессий	204
Оценка вероятностей	205
Оценка ранжирования	205
Использование данных для оценки	206
Независимые оценочные данные	206
Независимость на практике	207
Оценка для подгрупп населения	208
Приемлемый объем данных	210
Сравнение интеллектов	211
Рабочие точки	211
Кривые	212
Субъективные оценки	212
Изучение ошибок	213
Переосмысление опыта пользователя	214
Предсказание худшей ситуации	214
Итог главы	215
Темы для размышлений	216
Глава 20. Машинное обучение интеллекта	217
Как работает машинное обучение	217
Плюсы и минусы сложности	219
Недообучение	219
Переобучение	220
Поиск разумного компромисса	220
Конструирование признаков	221
Преобразование данных в удобную форму	221
Содействие модели в использовании данных	223
Нормализация значений признаков	224

Выявление скрытой информации.....	224
Расширение контекста	225
Устранение лишних признаков	225
Моделирование.....	226
Параметры сложности.....	226
Выявление переобучения	227
Итог главы	229
Темы для размышлений.....	229
Глава 21. Структурирование интеллекта	230
Причины структурирования интеллекта	230
Свойства правильно структурированного интеллекта.....	231
Способы структурирования интеллекта	232
Разделение признаков	232
Конкурентный поиск моделей.....	234
Распределение ошибок	235
Метамодель.....	235
Секвенированная модель.....	237
Разделение по контекстам.....	238
Замещение	239
Итог главы	240
Темы для размышлений.....	241
Часть V. ОРКЕСТРОВКА ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ.....	242
Глава 22. Понятие оркестровки интеллекта.....	243
Что такое хорошая оркестровка	244
Зачем нужна оркестровка	244
Изменение цели.....	245
Изменение пользователей	245
Изменение проблем	246
Изменение интеллекта.....	247
Изменение затрат.....	247
Злоупотребления	248
Команда оркестровки.....	248
Итог главы	249
Темы для размышлений.....	249
Глава 23. Среда оркестровки интеллекта	251
Мониторинг критериев успеха	251
Изучение взаимодействий.....	253
Оптимизация опыта.....	254
Переопределение интеллекта.....	255
Создание интеллекта.....	256

Критерии инвестирования в создание интеллекта.....	257
Итог главы	257
Темы для размышлений.....	258
Глава 24. Работа над ошибками	259
Худшее, что могло случиться	259
Причины выхода системы из строя.....	260
Отказы системы.....	261
Отказы модели.....	261
Ошибки интеллекта.....	262
Деградация интеллекта.....	262
Снижение количества ошибок.....	263
Инвестиции в развитие интеллекта.....	263
Настройка значимости опыта	264
Установка ограничителей	264
Переопределение ошибок.....	265
Итог главы	265
Темы для размышлений.....	266
Глава 25. Злоумышленники и злоупотребления.....	267
Злоупотребления – это бизнес.....	267
Масштабы злоупотреблений.....	268
Оценка вашего риска	269
Признаки злоупотреблений.....	270
Способы борьбы со злоупотреблениями	270
Увеличение стоимости продукта.....	270
Снижение привлекательности для злоумышленников	271
Машинное обучение против злоумышленников	271
Отключение злоумышленника от системы	271
Итог главы	272
Темы для размышлений.....	272
Глава 26. В шаге от собственной интеллектуальной системы	273
Контрольный список разработчика	273
Подход к проекту интеллектуальной системы.....	274
Планирование интеллектуального опыта	274
Планирование внедрения интеллектуальной системы.....	276
Подготовка к созданию интеллекта	278
Управляйте вашей интеллектуальной системой.....	279
Итог главы	280
Темы для размышлений.....	280
Предметный указатель.....	281