

УДК 004.89, 004.492
ББК 32.972
Ч58

Чио К., Фримэн Д.

Ч58 Машинное обучение и безопасность / пер. с англ. А. В. Снастина. – М.: ДМК Пресс, 2020. – 388 с.: ил.

ISBN 978-5-97060-713-8

Способна ли технология машинного обучения решить проблемы компьютерной безопасности? Или надежда на это является лишь следствием повышенного внимания к машинному обучению?

С помощью этой книги вы изучите способы применения машинного обучения в задачах обеспечения безопасности, таких как выявление вторжения извне, классификация вредоносных программ и анализ сетевой среды. Особое внимание уделено задачам по созданию работоспособных, надежных масштабируемых систем извлечения и анализа данных в сфере обеспечения безопасности.

Издание предназначено инженерам по обеспечению безопасности, а также специалистам по обработке данных научными методами.

УДК 004.89, 004.492
ББК 32.972

Original English language edition published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472. Copyright © 2018 Clarence Chio and David Freeman. Russian-language edition copyright © 2020 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 978-1-491-97990-7 (англ.)
ISBN 978-5-97060-713-8 (рус.)

Copyright © 2018 Clarence Chio and David Freeman
© Оформление, издание, перевод, ДМК Пресс, 2020

Содержание

Отзывы	5
Предисловие	11
Благодарности.....	15
 Глава 1. Машинное обучение и безопасность	 16
Общий обзор потенциальных киберугроз.....	18
Экономическая подоплека кибератак	21
Рынок услуг взломщиков	22
Косвенная монетизация.....	22
Подведем итоги	23
Что такое машинное обучение	24
Чем не является машинное обучение	25
Другие варианты использования машинного обучения	27
Практические варианты использования машинного обучения для обеспечения безопасности.....	27
Борьба со спамом: итеративный подход	30
Ограничения машинного обучения в сфере безопасности.....	40
 Глава 2. Классификация и кластеризация	 42
Машинное обучение: задачи и методики.....	42
Машинное обучение на практике: работающий пример	45
Тренировка алгоритмов машинного обучения	50
Семейства моделей.....	50
Функция потерь	53
Оптимизация	54
Алгоритмы классификации с учителем	57
Логистическая регрессия	57
Деревья решений	59
Леса деревьев решений	63
Метод опорных векторов	65
Наивный байесовский классификатор	67
Метод k ближайших соседей.....	70
Нейронные сети	71
Практические аспекты классификации	73
Выбор семейства моделей.....	73
Формирование процесса тренировки данных	74

Выбор признаков	78
Переподргонка и недоподргонка	79
Выбор пороговых значений и сравнение моделей	81
Кластеризация	82
Алгоритмы кластеризации	83
Оценка результатов кластеризации	93
Резюме	95

Глава 3. Выявление аномалий..... 97

Когда следует использовать методы выявления аномалий вместо обучения с учителем	98
Выявление вторжений с эвристиками	99
Методы, управляемые данными	101
Конструирование признаков для выявления аномалий.....	104
Выявление вторжения на хост.....	104
Выявление вторжения в сеть	107
Выявление вторжений в веб-приложение	111
Краткие итоги	112
Выявление аномалий с помощью данных и алгоритмов	113
Прогнозирование (машинное обучение с учителем)	114
Статистические метрики	125
Точность аппроксимации (качество подргонки)	126
Алгоритмы машинного обучения без учителя.....	132
Методы, основанные на плотностях.....	136
Краткие итоги	138
Трудности применения машинного обучения для выявления аномалий	139
Ответная реакция и ослабление воздействия	140
Практические аспекты проектирования систем.....	142
Оптимизация объяснимости	142
Удобство сопровождения систем выявления аномалий.....	143
Внедрение обратной связи с человеком	144
Снижение воздействий состязательности.....	144
Резюме.....	144

Глава 4. Анализ вредоносного программного обеспечения..... 145

Что такое вредоносное программное обеспечение	146
Классификация вредоносного программного обеспечения	148
Вредоносное программное обучение: что скрывается внутри.....	152
Генерация признаков	166
Сбор данных.....	167
Генерация признаков	169
Выбор признаков	193
От признаков к классификации	197
Как получить образцы и метки вредоносного программного обеспечения	200
Резюме.....	201

Глава 5. Анализ сетевого трафика	202
Теория защиты сетей.....	204
Управление доступом и аутентификация.....	204
Выявление вторжений	205
Обнаружение атакующих внутри сети.....	205
Защита, основанная на обработке данных	206
Приманка для злоумышленников	207
Резюме.....	207
Машинное обучение и обеспечение безопасности сети	207
От перехваченных данных к признакам	208
Угрозы в сетевой среде.....	213
Ботнет и защита от него.....	218
Создание модели прогнозирования для классификации сетевых атак	224
Исследование данных	226
Подготовка данных	230
Классификация	235
Обучение с учителем.....	237
Обучение с частичным привлечением учителя	243
Обучение без учителя.....	244
Расширенное ансамблирование	249
Резюме.....	254
 Глава 6. Защита потребительской веб-среды.....	 255
Монетизация в потребительской веб-среде	256
Типы мошенничества и данные, которые могут защитить.....	257
Аутентификация и перехват учетной записи.....	257
Создание учетной записи	264
Финансовое мошенничество	269
Деятельность ботов	272
Обучение с учителем для решения задач по выявлению нарушений	277
Метки для данных	278
Холодный запуск и горячий запуск.....	279
Ложноположительные и ложноотрицательные результаты	280
Несколько вариантов ответной реакции	281
Крупномасштабные атаки	281
Кластеризация нарушений	282
Пример: кластеризация доменов спама	283
Генерация кластеров	284
Оценка кластеров	289
Дальнейшие направления кластеризации	294
Резюме.....	295
 Глава 7. Производственные системы.....	 296
Определение зрелости и масштабируемости систем машинного обучения	296

Важные аспекты систем машинного обучения для обеспечения безопасности.....	297
Качество данных.....	298
Проблема: необъективность данных	298
Проблема: неточность меток.....	300
Решения: качество данных	300
Проблема: отсутствующие (потерянные) данные.....	302
Решения: отсутствующие (потерянные) данные	302
Качество модели.....	305
Проблема: оптимизация гиперпараметров	306
Решения: оптимизация гиперпараметров	307
Дополнительные функции: циклы обратной связи, А/В-тестирование моделей	311
Воспроизводимые и объяснимые результаты.....	315
Эффективность.....	319
Цель: минимальные задержки, высокая масштабируемость.....	319
Оптимизация эффективности.....	320
Горизонтальное масштабирование с помощью распределенных вычислительных программных сред	323
Использование облачных сервисов.....	328
Удобство сопровождения	330
Проблема: проверка контрольных точек, управление версиями и развертывание моделей.....	331
Цель: амортизация отказов	332
Цель: легкость настройки и конфигурации.....	333
Мониторинг и система оповещения	333
Безопасность и надежность	335
Функция: устойчивость и надежность работы во враждебных средах.....	335
Функция: защита и гарантии секретности данных	336
Обратная связь и удобство использования	337
Резюме.....	338

Глава 8. Состязательное машинное обучение

Терминология	340
Важность состязательного машинного обучения	341
Опасные уязвимости в алгоритмах машинного обучения.....	342
Мобильность атак.....	345
Методика атак: заражение модели	346
Пример: заражающая атака на бинарный классификатор.....	349
Знания атакующего.....	355
Защита от заражающих атак.....	356
Методика атаки: искажающая атака	358
Пример: искажающая атака на бинарный классификатор	359
Защита от искажающих атак	364
Резюме.....	365

Приложение А. Дополнительный материал к главе 2	367
Подробнее о метриках	367
Размер моделей логистической регрессии	368
Реализация функции стоимости для метода логистической регрессии	368
Минимизация функции стоимости.....	369
 Приложение Б. Разведка на основе открытых	
источников	374
Материалы разведки для обеспечения безопасности	374
Геолокация	376
 Предметный указатель.....	377