

3281 СС худ
8812 худ
120

Предиктивное моделирование

на практике

Изучение предиктивных методов моделирования в контексте применения их для решения практических задач показывает, что эти методы обладают рядом преимуществ перед традиционными. Важнейшими из них являются: высокая точность и надежность, возможность учета нелинейных и стохастических зависимостей, а также возможность определения физических законов, лежащих в основе явлений. Применение предиктивных методов в различных отраслях промышленности и строительства позволяет решать сложные инженерные задачи с высокой точностью и эффективностью.

Макс Кун и Кэлл Джонсон

3281 СС худ
Ростехнадзор

ПИТЕР®
Санкт-Петербург · Москва · Екатеринбург · Воронеж
Нижний Новгород · Ростов-на-Дону
Самара · Минск

2019

ББК 22.183.5

УДК 519.8

К91

Кун Макс, Джонсон Кьюлл

- К91 Предиктивное моделирование на практике. — СПб.: Питер, 2019. — 640 с.: ил. — (Серия «Для профессионалов»).

ISBN 978-5-4461-1039-1

«Предиктивное моделирование на практике» охватывает все аспекты прогнозирования, начиная с ключевых этапов предварительной обработки данных, разбиения данных и основных принципов настройки модели. Все этапы моделирования рассматриваются на практических примерах из реальной жизни, в каждой главе дается подробный код на языке R.

Эта книга может использоваться как введение в предиктивные модели и руководство по их применению. Читатели, не обладающие математической подготовкой, оценят интуитивно понятные объяснения конкретных методов, а внимание, уделяемое решению актуальных задач с реальными данными, поможет специалистам, желающим повысить свою квалификацию.

Авторы постарались избежать сложных формул, для усвоения основного материала достаточно понимания основных статистических концепций, таких как корреляция и линейный регрессионный анализ, но для изучения углубленных тем понадобится математическая подготовка.

Для работы с книгой нужно иметь базовые знания о языке R.

Премия Ziegel Technometrics 2014.

16+ (В соответствии с Федеральным законом от 29 декабря 2010 г. № 436-ФЗ.)

ББК 22.183.5

УДК 519.8

Права на издание получены по соглашению с Springer Science+Business Media, LLC. Все права защищены. Никакая часть данной книги не может быть воспроизведена в какой бы то ни было форме без письменного разрешения владельцев авторских прав.

Информация, содержащаяся в данной книге, получена из источников, рассматриваемых издательством как надежные. Тем не менее, имея в виду возможные человеческие или технические ошибки, издательство не может гарантировать абсолютную точность и полноту приводимых сведений и не несет ответственности за возможные ошибки, связанные с использованием книги. Издательство не несет ответственности за доступность материалов, ссылки на которые вы можете найти в этой книге. На момент подготовки книги к изданию все ссылки на интернет-ресурсы были действующими.

ISBN 978-1461468486 англ.

First published in English under the title Applied Predictive Modeling, 1st Edition by Max Kuhn, Kjell Johnson by Springer Science+Business Media, LLC
© 2013, Springer Science+Business Media New York

This edition has been translated and published under licence from Springer Science+Business Media, LLC.

Springer Science+Business Media, LLC takes no responsibility and shall not be made liable for the accuracy of the translation.

ISBN 978-5-4461-1039-1

© Перевод на русский язык ООО Издательство «Питер», 2019

© Издание на русском языке, оформление ООО Издательство «Питер», 2019

© Серия «Для профессионалов», 2019

Краткое содержание

Предисловие	18
Глава 1. Введение	21

ЧАСТЬ I ОБЩИЕ СТРАТЕГИИ

Глава 2. Краткий обзор процесса предиктивного моделирования.....	42
Глава 3. Предварительная обработка данных.....	51
Глава 4. Переобучение и настройка модели	87

ЧАСТЬ II РЕГРЕССИОННЫЕ МОДЕЛИ

Глава 5. Измерение эффективности регрессионных моделей	122
Глава 6. Модели с признаками линейной регрессии	128
Глава 7. Нелинейные регрессионные модели	169
Глава 8. Древовидные модели. Модели на базе правил.....	202
Глава 9. Обзор моделей растворимости.....	254
Глава 10. Практический пример: сопротивление сжатию бетонных смесей.....	257

ЧАСТЬ III КЛАССИФИКАЦИОННЫЕ МОДЕЛИ

Глава 11. Определение эффективности в классификационных моделях.....	278
Глава 12. Дискриминантный анализ и другие линейные классификационные модели	307
Глава 13. Нелинейные классификационные модели	365
Глава 14. Деревья классификации и модели на базе правил.....	407
Глава 15. Сравнительный анализ моделей для заявок на получение грантов.....	456
Глава 16. Решение проблемы дисбаланса классов	460
Глава 17. Практикум: планирование заданий.....	488

ЧАСТЬ IV ПРОЧИЕ ВОПРОСЫ ПРЕДИКТИВНОГО МОДЕЛИРОВАНИЯ

Глава 18. Определение важности предикторов	506
Глава 19. Выбор признаков.....	531
Глава 20. Факторы, влияющие на эффективность модели.....	567

ПРИЛОЖЕНИЯ

Приложение А. Краткая сводка различных моделей	596
Приложение Б. Введение в R.....	599
Приложение В. Рекомендуемые веб-сайты.....	616
Список источников	619