

УДК 378.147:620.9(075)  
ББК Ч481.253.1:31.19я  
Ц26

*Печатается по решению редакционно-издательского совета  
Казанского национального исследовательского технологического университета*

*Рецензенты:*  
*д-р техн. наук, проф. Г. В. Вагапов*  
*канд. техн. наук, доц. Т. Х. Мухаметгалеев*

**Цвенгер И. Г.**

**Ц26** Машинное обучение в энергоснабжении : учебное пособие / И. Г. Цвенгер; Минобрнауки России, Казан. нац. исслед. технол. ун-т. — Казань : Изд-во КНИТУ, 2024. — 112 с.

ISBN 978-5-7882-3489-2

Содержит основные сведения о машинном обучении применительно к вопросам энергоэффективности зданий мегаполиса. Рассмотрены вопросы обработки входных данных, обучения модели и интерпретации результатов ее работы, а также примеры практического применения методов машинного обучения для закрепления теоретических знаний.

Предназначено для обучающихся направления 13.04.02 «Электроэнергетика и электротехника» (профиль «Цифровизация электротехнических комплексов и систем нефтегазохимических производств»).

Подготовлено на кафедре электропривода и электротехники.

**УДК 378.147:620.9(075)**  
**ББК Ч481.253.1:31.19я**

ISBN 978-5-7882-3489-2

© Цвенгер И. Г., 2024

© Казанский национальный исследовательский  
технологический университет, 2024

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ .....	4
1. РАБОТА С ДАННЫМИ .....	5
1.1. Среда разработки и этапы исследования .....	5
1.2. Загрузка данных .....	6
1.3. Первичный анализ данных .....	8
1.4. Интенсивность использования энергии .....	10
1.5. Поиск зависимостей .....	13
1.6. Отсутствующие и аномальные данные .....	19
1.7. Разведочный анализ данных .....	27
1.8. Поиск взаимосвязей .....	31
1.9. Конструирование и выбор признаков .....	36
1.10. Выбор базового уровня и разделение данных .....	48
2. ОБУЧЕНИЕ МОДЕЛИ .....	51
2.1. Заполнение отсутствующих значений .....	51
2.2. Масштабирование признаков (нормализация) .....	52
2.3. Оценка и выбор модели .....	54
2.4. Градиентный бустинг (Gradient Boosting) .....	59
2.5. Библиотеки, реализующие градиентный бустинг .....	61
2.6. Описание класса GradientBoostingRegressor .....	62
2.7. Определение гиперпараметров .....	67
2.8. Настройка гиперпараметров .....	68
2.9. Описание процесса перекрестной проверки (Cross Validation) .....	70
2.10. Описание класса RandomizedSearchCV .....	71
2.11. Анализ результатов настройки гиперпараметров .....	74
2.12. Описание класса GridSearchCV .....	76
2.13. Зависимость точности модели от числа деревьев .....	79
2.14. Глубина деревьев в градиентном бустинге .....	81
2.15. Минимальное число объектов в листе .....	82
2.16. Оценка точности и производительности модели .....	84
3. ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ РАБОТЫ МОДЕЛИ .....	94
3.1. Загрузка данных для обучения модели .....	95
3.2. Интерпретация модели .....	97
3.3. Анализ важности признаков .....	102
3.4. Анализ ошибок в прогнозе .....	104
ЗАКЛЮЧЕНИЕ .....	108
ЛИТЕРАТУРА .....	110