

УДК 004.8

Рекомендовано к изданию методическим советом ПГУТИ, протокол №41, от 27.02.2017 г.

Рецензент:

научный сотрудник, к.э.н., доцент ИПУСС РАН

Моисеева Т.В.

Пальмов, С. В.

- П** Интеллектуальный анализ данных: учебное пособие / С. В. Пальмов. – Самара: ПГУТИ, 2017. – 124 с.

Аннотация

Рассматриваются основы методов искусственного интеллекта, история возникновения области искусственного интеллекта и становления её как науки. Рассмотрены базовые алгоритмы (ассоциативные правила, деревья решений, нейронные сети и т.д.), позволяющие решать типичные задачи для вышеуказанной области: представление знаний и поиск.

Предназначено в качестве учебного пособия для студентов направления подготовки 09.04.02. «Информационные системы и технологии» по дисциплине «Интеллектуальный анализ данных».

ISBN

©, Пальмов С.В., 2017

Оглавление

Глава 1. Основные понятия интеллектуального анализа данных.....	6
1.1. Определение интеллектуального анализа данных	6
1.2. Некоторые бизнес-приложений ИАД.....	9
1.3. Классы систем Data Mining	14
1.4. Общие особенности ИАД.....	19
Вопросы для самоконтроля:	23
Глава 2. Обзор возможностей генетических алгоритмов.....	24
2.1. Эволюционные алгоритмы.....	24
2.2. Генетические алгоритмы	25
Вопросы для самоконтроля:	27
Глава 3. Извлечение знаний из памяти эксперта.....	28
3.1. Процедура взаимодействия инженера по знаниям с экспертом	28
3.2. Классификация методов работы с экспертами	28
3.3. Пассивные методы	30
3.4. Активные индивидуальные методы	33
3.5. Активные групповые методы.....	40
3.6. Экспертные игры.....	43
Вопросы для самоконтроля:	47
Глава 4. Структурирование знаний	48
4.1. Система понятий	48
4.2. Семантические отношения.....	50
4.3. Стратегии принятия решений	54
Вопросы для самоконтроля:	62
Глава 5. Случайный лес.....	63
5.1. Что такое случайный лес?	63
5.2. Случайный лес. Прогнозирование	67
5.3. Случайный лес и сегментация	69
5.4. Пропущенные значения.....	71
5.5. Значимость переменных	72
5.6. Случайный лес: достоинства и недостатки.....	74
5.7. Пример использования случайного леса	75
Вопросы для самоконтроля:	76
Глава 6. Краткий обзор алгоритма k -средних (k -means).....	77
6.1. Общие сведения.....	77
Вопросы для самоконтроля:	80

Глава 7. Байесовская классификация. Краткий обзор возможностей	81
Вопросы для самоконтроля:	83
Глава 8. Метод опорных векторов.....	84
8.1. Общие сведения.....	84
8.2. Линейный SVM	86
Вопросы для самоконтроля:	88
Глава 9. Метод «ближайшего соседа» или системы рассуждений на основе аналогичных случаев. Краткий обзор возможностей.....	89
Вопросы для самоконтроля:	91
Глава 10. Введение в Social Mining	92
10.1. Понятие социальной сети и её анализа	92
10.2. Виртуальная социальная сеть.....	94
10.3. Задачи Social Mining	96
Вопросы для самоконтроля:	98
Глава 11. Введение в Web Mining.....	99
11.1. Основные понятия и принципы	99
11.2. Сложности анализа данных из сети Интернет.....	100
11.3. Этапы Web Mining.....	100
11.4. Категории Web Mining.....	101
Вопросы для самоконтроля:	103
Глава 12. Обобщённые ассоциативные правила	104
12.1. Введение.....	104
12.2. Описание задачи.....	105
12.3. Определение «интересных» правил	106
12.4. Алгоритм вычисления обобщенных ассоциативных правил	109
12.5. Базовый алгоритм поиска часто встречающихся множеств	110
12.6. Улучшенный алгоритм поиска часто встречающихся множеств	112
Вопросы для самоконтроля:	114
Глава 13. Алгоритм поиска ассоциативных правил FPG	115
13.1. Введение.....	115
13.2. Описание алгоритма	115
Вопросы для самоконтроля:	124

Глоссарий.....	125
Список литературы	126
Заключение	127