

[004.6+519.254](076.5)

УДК 519.254(076)

ББК 32.973.3я73

С32

Рецензенты:

кафедра интеллектуальных систем
Белорусского национального технического университета
(протокол №4 от 15.11.2018);

старший научный сотрудник государственного научного учреждения
«Объединенный институт проблем информатики Национальной академии наук
Беларуси» кандидат физико-математических наук С. В. Чебаков

Серебряная, Л. В.

С32 Модели и методы обработки данных в информационных системах :
учеб.-метод. пособие / Л. В. Серебряная, В. В. Потараев, Е. П. Фадеева. –
Минск : БГУИР, 2019. – 67 с. : ил.
ISBN 978-985-543-506-9.

Рассмотрены различные типы и особенности функционирования современных информационных систем. Приведены модели и методы работы с данными, используемые в информационных системах. Предложены шесть лабораторных работ, реализация которых позволит практически закрепить знания по курсу «Модели и методы обработки данных в информационных системах».

УДК 519.254(076)

ББК 32.973.3я73

ISBN 978-985-543-506-9



- © Серебряная Л. В., Потараев В. В., Фадеева Е. П., 2019
- © УО «Белорусский государственный университет информатики и радиоэлектроники», 2019

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	4
ТЕОРЕТИЧЕСКАЯ ЧАСТЬ	
1 СТРУКТУРА И СОСТАВ ИНФОРМАЦИОННОЙ СИСТЕМЫ	6
2 СИСТЕМЫ КЛАССИФИКАЦИИ ДАННЫХ	8
3 СИСТЕМЫ КОДИРОВАНИЯ	10
4 ИНТЕЛЛЕКТУАЛЬНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ	12
5 ГЕНЕТИЧЕСКИЙ АЛГОРИТМ	15
5.1 Кодирование информации и формирование популяции	16
5.2 Селекция	18
5.3 Скрещивание и формирование нового поколения	19
5.4 Мутация	21
5.5 Настройка параметров генетического алгоритма	21
ЛАБОРАТОРНЫЙ ПРАКТИКУМ	
ЛАБОРАТОРНАЯ РАБОТА №1	
ПРИМЕНЕНИЕ ЗАКОНОВ ЗИПФА ДЛЯ ОБРАБОТКИ ТЕКСТОВОЙ	
ИНФОРМАЦИИ	24
1.1 Первый закон Зипфа «ранг – частота»	25
1.2 Второй закон Зипфа «количество – частота»	25
1.3 Весовые коэффициенты	27
ЛАБОРАТОРНАЯ РАБОТА №2	
ПОИСК ТЕКСТОВОЙ ИНФОРМАЦИИ ПО ЗАДАННОМУ НАБОРУ	
КЛЮЧЕВЫХ СЛОВ	28
2.1 Особенности поиска информации	28
2.2 Алгоритм поиска	31
ЛАБОРАТОРНАЯ РАБОТА №3	
АВТОМАТИЧЕСКАЯ РУБРИКАЦИЯ ТЕКСТОВОЙ ИНФОРМАЦИИ	
ПО ОБРАЗЦУ	32
3.1 Алгоритм автоматической рубрикации текстов по образцу	32
3.2 Разделение объектов на N классов методом персептрона	35
ЛАБОРАТОРНАЯ РАБОТА №4	
МЕТОД АВТОМАТИЧЕСКОГО РЕФЕРИРОВАНИЯ ТЕКСТОВОЙ	
ИНФОРМАЦИИ	41
4.1 Автоматическое реферирование и аннотирование текстов	41
ЛАБОРАТОРНАЯ РАБОТА №5	
РЕАЛИЗАЦИЯ ГЕНЕТИЧЕСКОГО АЛГОРИТМА	43
5.1 Канонический генетический алгоритм	44
ЛАБОРАТОРНАЯ РАБОТА №6	
ПРИМЕНЕНИЕ НЕЙРОННОЙ И СЕМАНТИЧЕСКОЙ СЕТЕЙ	
ДЛЯ ОБРАБОТКИ ТЕКСТОВОЙ ИНФОРМАЦИИ	47
6.1 Персептрон и сеть Хопфилда	48
6.2 Применение нейронной сети Хопфилда для ответа на вопрос	53
6.3 Поиск ответа на вопрос с помощью семантической сети	59
ЗАКЛЮЧЕНИЕ	65
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	66