

УДК 004.056.5(075.8)

ББК 32.97я73

Ц349

Печатается по решению кафедры информационно-аналитических систем безопасности Института компьютерных технологий и информационной безопасности Южного федерального университета (протокол № 10 от 10 июня 2022 г.)

Рецензенты:

доктор технических наук, профессор, заведующий кафедрой информатики Таганрогского института имени А. П. Чехова (филиал) РГЭУ (РИНХ)

Я. Е. Ромм

доктор технических наук, профессор,
профессор кафедры информационно-аналитических систем безопасности Южного федерального университета *С. Л. Беляков*

Целых, А. Н.

Ц349 Принятие решений на основе методов машинного обучения : учебное пособие по курсам «Модели и методы инженерии знаний», «Методы анализа больших данных» / А. Н. Целых, Н. В. Драгныш, Э. М. Котов ; Южный федеральный университет. – Ростов-на-Дону ; Таганрог : Издательство Южного федерального университета, 2022. – 113 с.

ISBN 978-5-9275-4246-8

Пособие посвящено рассмотрению подходов к разработке экспертной системы принятия решений для медицинских целей, с использованием больших данных для формирования требований к системе самодиагностики и оказания первой помощи. Практический аспект включает проектирование структуры БД диагнозов и разработку алгоритмы взаимодействия с ней, а так же разработку прототипа интерфейса администратора системы.

Пособие предназначено для студентов высших учебных заведений, обучающихся по специальности 10.05.04 Информационно-аналитические системы безопасности (специализация: "Автоматизация информационно-аналитической деятельности") по курсу "Математические методы анализа больших данных" и направлению 10.03.01 Информационная безопасность (направленность: "Информационно-аналитические системы безопасности") по курсу "Модели и методы инженерии знаний".

УДК 004.056.5(075.8)

ББК 32.97я73

ISBN 978-5-9275-4246-8

© Южный федеральный университет, 2022

© Целых А. Н., Драгныш Н. В., Котов Э. М., 2022

© Оформление. Макет. Издательство

Южного федерального университета, 2022



СОДЕРЖАНИЕ

ВВЕДЕНИЕ	6
1. МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ	8
1.1. Описание предметной области	8
1.2. Классификация	9
1.3. Кластеризация	11
1.4. Машинное обучение в задаче извлечения информации	12
1.5. Модели машинного обучения	13
1.5.1. Наивный байесовский классификатор	13
1.5.2. Метод опорных векторов	14
1.5.3. Нейронная сеть	14
1.6. Выбор метрик оценки машинного обучения	15
1.6.1. Матрица ошибок	15
1.6.2. Точность	16
1.6.3. Recall	16
1.6.4. Precision	16
1.6.5. F1 score	17
2. БОЛЬШИЕ ДАННЫЕ В СФЕРЕ САМОДИАГНОСТИКИ	18
2.1. Анализ сферы самодиагностики	18
2.1.1. Преимущества самодиагностики	18
2.1.2. Недостатки и ограничение самодиагностики	19
2.1.3. Сравнение имеющихся систем самодиагностики	20
2.2. Анализ области оказания первой помощи	23
2.2.1. Актуальность оказания первой помощи	23
2.2.2. Проблематика оказания первой помощи	24
2.2.3. Использование информационных технологий при оказании первой помощи	25
2.2.4. Сравнение существующих систем оказания первой помощи	25

Содержание

2.3. Основания для разработки СС и ОПП	28
2.4. Проектирование структуры разрабатываемой системы	29
3. РАЗРАБОТКА МОДЕЛИ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ	31
3.1. Логическая декомпозиция предметной области	31
3.1.1. Исходные знания	31
3.1.2. Классификация и извлечение знаний	34
3.2. Разработка фреймовой модели	39
3.3. Формирование обобщенного алгоритма работы механизма логического вывода	40
4. ПРОЕКТИРОВАНИЕ СЕРВЕРНОЙ ЧАСТИ СИСТЕМЫ	43
4.1. Общие сведения о системе	43
4.2. Архитектура серверной части системы	43
4.2.1. Определение вариантов использования	43
4.2.2. Структура серверной части СС и ОПП	48
4.2.3. Размещение программных компонентов системы СС и ОПП	51
4.3. Выбор модели жизненного цикла системы и средств реали- зации программных компонентов	53
4.3.1. Выбор технологий реализации интерфейса администра- тора	54
4.3.2. Выбор технологии реализации веб-серверов	56
4.3.3. Выбор СУБД	56
4.3.4. Выбор прокси-сервера	57
4.4. Проектирование БД-диагнозов	57
4.4.1. Концептуальная модель предметной области	57
4.4.2. Инфологическая модель базы данных	59
4.4.3. Получение метаинформации о базе данных	61
4.4.4. Даталогическая структура базы данных	63
4.5. Построение объектной модели базы данных	67
4.5.1. Основной алгоритм построения объектной модели	67
4.5.2. Разделение таблиц на основные и промежуточные	68
4.5.3. Разделение полей данных и внешних ключей	70
4.5.4. Нахождение связей «много ко многим»	72
4.6. Разработка графического интерфейса администратора	73

Содержание

4.6.1. Определение состояний интерфейса и условий пере-хода	73
4.6.2. Прототипирование интерфейса администратора	75
4.7. Разработка протокола взаимодействия клиентской и сервер- ной частей СС и ОПП	77
4.7.1. Выбор архитектуры API	77
4.7.2. Проектирование API	78
5. ПРОЕКТИРОВАНИЕ КЛИЕНТСКОЙ ЧАСТИ СИСТЕМЫ ...	82
5.1. Выбор архитектуры и типа пользовательского интерфейса	82
5.2. Выбор подхода, языка программирования и среды разра- ботки	82
5.3. Разработка диаграммы вариантов использования	83
5.4. Проектирование структурной схемы СС и ОПП	88
5.5. Проектирование компонентов СС и ОПП	90
5.5.1. Проектирование интерфейса пользователя	90
5.5.2. Разработка алгоритма работы механизма логического вывода	91
5.6. Проектирование классов приложения СС и ОПП	97
6. РАЗМЕЩЕНИЕ И ТЕСТИРОВАНИЕ КОМПОНЕНТОВ СИСТЕМЫ	100
6.1. Программно-техническая среда функционирования серверной части	100
6.2. Тестирование СС и ОПП	101
6.2.1. Функциональное тестирование серверной части СС и ОПП.	101
6.2.2. Оценочное тестирование серверной части СС и ОПП	103
ЗАКЛЮЧЕНИЕ	109
СПИСОК ЛИТЕРАТУРЫ	110