

Міністерство освіти і науки України

Литвин В. В.

Методи та засоби інженерії даних та знань

*Рекомендовано як навчальний посібник
для студентів вищих навчальних закладів*

**Видавництво ПП “Магнолія 2006”
Львів 2024**

УДК 004.415(075.8)
ББК 32.973.25я73
Л 64

Відтворення цієї книги або будь-якої її частини заборонено без письмової згоди видавництва. Будь-які спроби порушення авторських прав будуть переслідуватися у судовому порядку.

Гриф надано Міністерством освіти і науки України

Рецензенти:

Г. Г. Цегелик – доктор фізико-математичних наук, професор, завідувач кафедри математичного моделювання соціально-економічних процесів Львівського Національного університету імені Франка;

Я. М. Матвійчук – доктор технічних наук, професор, завідувач кафедри інформаційних систем та технологій інституту підприємництва та перспективних технологій при Національному університеті „Львівська політехніка”;

Б. П. Русин – доктор технічних наук, завідувач відділу ФМІ НАН України імені Карпенка

Литвин В. В.

Л 64 **Методи та засоби інженерії даних та знань:** Навч. посіб. – Львів: “Магнолія 2006”, 2024. – 241 с.

ISBN 978-617-574-044-6

“Магнолія 2006”

Викладено основні методи інженерії даних та знань, особливу увагу звернуто на технології видобування, нагромадження, структуризації, формування, опрацювання даних та знань. Висвітлено особливості технологій інженерії даних та знань, їх теоретичні та прикладні аспекти. Розглянуто моделі онтологічних систем та методику розроблення онтологій. Детально описано побудову онтології за допомогою програмного засобу Protégé. Розглянуто методи машинного навчання, які використовуються під час інженерії даних та знань.

Навчальний посібник призначений для студентів, що навчаються за напрямками підготовки „Комп’ютерні науки”, „Системний аналіз”, для магістрів спеціальностей, які базуються на цих напрямках підготовки, а також для магістрів спеціальності “Консолідована інформація”.

УДК 004.415(075.8)
ББК 32.973.25я73

© Литвин В.В., 2024
© “Магнолія 2006”, 2024

ISBN 978-617-574-044-6

ЗМІСТ

РОЗДІЛ 1. Основні поняття інженерії даних та знань.....	9
1.1. Методи придбання знань.....	10
1.1.1. Категорії навчання.....	11
1.1.2. Придбання знань на метарівні.....	11
1.1.3. Придбання знань із прикладів.....	12
1.2. Базові поняття менеджменту знань.....	13
1.2.1. Видобування знань.....	14
1.2.2. Системи пізнання.....	15
1.2.3. Організація доступу до знань.....	15
1.2.4. Інновації в області автоматизації.....	16
1.2.5. Менеджмент знань та інформація.....	18
1.3. Онтологічний інжиніринг.....	18
1.3.1. Системи керування знаннями.....	19
1.3.2. Онтологія.....	21
Контрольні запитання.....	22
РОЗДІЛ 2. Теоретичні аспекти інженерії даних та знань.....	23
2.1. Поле знань.....	23
2.1.1. Мова опису поля знань.....	23
2.1.2. Семіотична модель поля знань.....	25
2.1.3. Мова опису поля знань.....	28
2.2. Стратегії одержання знань.....	28
2.3. Теоретичні аспекти видобування знань.....	32
2.3.1. Психологічний аспект.....	32
2.3.2. Лінгвістичний аспект.....	39
2.3.3. Гносеологічний аспект видобування знань.....	43
2.4. Теоретичні аспекти структурування знань.....	48
2.4.1. Історична довідка.....	48
2.4.2. Ієрархічний підхід.....	49
2.4.3. Традиційні методології структуризації.....	50
2.4.4. Об'єктно-структурний підхід.....	50
Контрольні запитання.....	54
РОЗДІЛ 3. Технології інженерії даних та знань.....	56
3.1. Класифікація методів практичного видобування знань.....	56
3.2. Комунікативні методи.....	59
3.2.1. Пасивні методи.....	59
3.2.2. Активні індивідуальні методи.....	62
3.2.3. Активні групові методи.....	70
3.3. Текстологічні методи.....	75

3.3.1. Методи структурування.....	79
3.3.2. Еволюція систем одержання знань.....	82
Контрольні запитання.....	86
РОЗДІЛ 4. Прикладні аспекти інженерії даних та знань.....	87
4.1. Латентні структури знань і психосемантика.....	87
4.1.1. Семантичні простори і психологічне градування.....	87
4.1.2. Методи багатовимірного градування.....	92
4.1.3. Використання метафор для виявлення структур знань.....	94
4.2. Метод репертуарних решіток.....	98
4.2.1. Основні поняття.....	98
4.2.2. Методи виявлення конструктів. Метод мінімального контексту.....	100
4.2.3. Аналіз репертуарних решіток.....	101
4.2.4. Автоматизовані методи.....	103
4.3. Керування знаннями.....	104
4.3.1. Що таке “керування знаннями”?.....	104
4.3.2. Структура менеджменту знань.....	107
4.3.3. Формування знань.....	108
4.3.4. Адміністрування.....	110
4.3.5. Мотивація.....	110
4.3.6. Особливості впровадження МЗ.....	111
4.3.7. Менеджмент знань.....	112
4.3.8. Керування знаннями і корпоративна пам’ять.....	116
4.3.9. Системи OMIS.....	118
4.3.10. Особливості розроблення OMIS.....	120
Контрольні запитання.....	121
РОЗДІЛ 5. Класифікація даних та знань.....	123
5.1. Важливість правильної класифікації.....	123
5.1.1. Класифікація й об’єктно-орієнтовне проектування.....	123
5.1.2. Труднощі класифікації.....	124
5.2. Ідентифікація класів і об’єктів.....	126
5.2.1. Класичний і сучасний підходи.....	126
5.2.2. Об’єктно-орієнтований аналіз.....	129
5.3. Ключові абстракції й механізми.....	133
5.3.1. Ключові абстракції.....	133
5.3.2. Ідентифікація механізмів.....	135
Висновки.....	136
Контрольні запитання.....	136
РОЗДІЛ 6. Онтології й онтологічні системи.....	137
6.1. Поняття онтології.....	137
6.2. Моделі онтології й онтологічної системи.....	145

6.3. Методології створення і “життєвий цикл” онтологій.....	150
6.4. Мови опису онтологій.....	150
Контрольні запитання.....	154
РОЗДІЛ 7. Програмні засоби побудови онтологій.....	155
7.1. Онтологія як засіб формалізації та алгоритмізації знань.....	155
7.1.1. Аналіз підходів до навчання онтологій.....	156
7.1.2. Загальні принципи проектування онтологій.....	158
7.1.3. Формати та стандарти подання інформації.....	159
7.1.4. Засоби для створення онтологій.....	162
7.2. Технологія розроблення онтологій в редакторі Protégé.....	163
7.2.1. Еволюція Protégé.....	163
7.2.2. Protégé-OWL. Мова Web онтологій OWL.....	165
7.2.3. Основні терміни та поняття у Protégé-OWL.....	167
7.2.4. Методика розроблення онтологій засобами Protégé.....	168
7.2.5. Створення онтологій.....	171
Контрольні запитання.....	183
РОЗДІЛ 8. Дерева прийняття рішень.....	184
8.1. Поняття дерева рішень.....	184
8.2. Структура дерева прийняття рішень.....	185
8.3. Ентропія і приріст інформації.....	186
8.4. Дерева прийняття рішень і булеві функції.....	189
8.5. ID3 алгоритм.....	190
8.6. Проблеми з критерієм приросту інформації. Gain Ratio і GINI Index.....	191
8.7. Оверфіттинг.....	192
Контрольні запитання.....	192
РОЗДІЛ 9. Нейронні мережі.....	193
9.1. Вступ.....	193
9.2. Перцептрон.....	194
9.3. Навчання перцептрона.....	195
9.4. Метод градієнтного спуску.....	199
9.5. Нелінійні перцептрони. Сигмоїд.....	200
9.6. Алгоритм зворотного поширення помилки в нейронних мережах.....	202
Контрольні запитання.....	204
РОЗДІЛ 10. Генетичні алгоритми.....	205
10.1. Вступ.....	205
10.2. Схема генетичного алгоритму.....	206
10.3. Генетичні операції.....	207
10.4. Подання даних.....	208
10.5. Відбір.....	210
Контрольні запитання.....	214
РОЗДІЛ 11. Байєсівське навчання.....	215

11.1. Вступ.....	215
11.2. Основні поняття теорії ймовірностей.....	215
11.3. Теорема Байєса, дані і гіпотези.....	216
11.4. МАР і задачі класифікації.....	217
11.5. Оптимальний байєсівський класифікатор і класифікатор Гіббса... ..	218
11.6. Наївний байєсівський класифікатор.....	220
11.7. Байєсівське навчання і нейронні мережі.....	221
Контрольні запитання.....	222
РОЗДІЛ 12. Стимульоване навчання.....	223
12.1. Вступ і постановка задачі.....	223
12.2. Оцінка поведінки агента.....	224
12.3. Багаторукі бандити.....	226
12.4. Доказово оптимальні алгоритми.....	226
12.5. Евристичні стратегії.....	229
12.6. Модель агентів з декількома станами.....	230
12.7. Пошук оптимальних стратегій у відомій моделі.....	231
Контрольні запитання.....	233
РОЗДІЛ 13. Асоціативні правила.....	234
13.1. Поняття асоціативного правила.....	234
13.2. Алгоритм Аргіогі побудови асоціативних правил.....	237
13.3. Сфера застосування асоціативних правил.....	240
Контрольні питання.....	240
Використана література.....	241

НАВЧАЛЬНЕ ВИДАННЯ

ЛИТВИН Василь Володимирович

Методи та засоби інженерії даних та знань

Навчальний посібник

Формат 70x100/16. Папір друк. №2.
Гарнітура Тип Таймс. Умовн. друк. арк. 14,93.

ПП “Магнолія 2006”
м. Львів-53, 79053, Україна,
тел.+350503701957
e-mail: magnol06@ukr.net

Свідоцтво про внесення суб’єкта видавничої справи
до Державного реєстру видавців, виготівників і
розповсюджувачів видавничої
продукції: серія ДК № 2534 від 21.06.2006 року,
видане Державним комітетом інформаційної політики,
телебачення та радіомовлення України

Надруковано у друкарні видавництва “Магнолія 2006”