

Міністерство освіти і науки України
Національна Академія наук України
Академія технологічних наук України
Інженерна академія України
Університет Гліндор, м. Рексхем, Великобританія
Технічний університет Лодзі, Польща
Технічний університет м. Рига, Латвія
Технологічний університет м. Таллінн, Естонія
Університет Екстрамадура, м. Бадахос, Іспанія
Гомельський державний університет ім. Ф. Скорини, Білорусь
Інститут проблем математичних машин і систем (ІПММС) НАН України
Інститут прикладної математики ім. М.В. Келдиша РАН, м. Москва, Росія
Національний технічний університет України
«Київський політехнічний інститут»
Полтавський національний технічний університет імені Юрія Кондратюка
Чернігівський національний технологічний університет

**МАТЕМАТИЧНЕ ТА ІМІТАЦІЙНЕ
МОДЕЛЮВАННЯ СИСТЕМ
МОДС 2017**

**ДВНАДЦЯТА МІЖНАРОДНА НАУКОВО-ПРАКТИЧНА
КОНФЕРЕНЦІЯ**

26 - 29 червня 2017 р.

Тези доповідей



Чернігів 2017

УДК 004.94(063)

ББК 22.193(06)

М34

Рекомендовано до друку вченою радою Чернігівського національного технологічного університету (протокол № 6 від 29.05.2017р.).

Редакційна колегія:

Казимир В.В., д-р техн. наук, професор, ЧНТУ

Литвинов В. В., д-р техн. наук, професор, ЧНТУ

Войцеховська М.М., аспірант, ЧНТУ

Житник О.Е., аспірант, ЧНТУ

Нехай В.В., аспірант, ЧНТУ

М34 **Математичне та імітаційне моделювання систем. МОДС 2017** : тези доповідей Дванадцятої міжнародної науково-практичної конференції (Чернігів, 26–29 червня 2017 р.) / М-во освіти і науки України, Нац. акад. наук України, Академія технологічних наук України, Інженерна академія України та ін. – Чернігів : ЧНТУ, 2017. – 444 с.

ISBN 978-966-2188-89-9

У збірник включені тези доповідей, які були представлені на конференції “Математичне та імітаційне моделювання систем. МОДС 2017”. У доповідях розглянуті наукові та методичні питання з напрямку моделювання складних екологічних, технічних, фізичних, економічних, виробничих, організаційних та інформаційних систем з використанням математичних та імітаційних методів.

УДК 004.94(063)

ББК 22.193(06)

ISBN 978-966-2188-89-9

© Чернігівський національний технологічний університет

В.В. Казимир, А.Ю. Солдатов ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ВЗАЄМОДІЇ МОБІЛЬНИХ ДОДАТКІВ З ВІРТУАЛЬНИМ СЕРЕДОВИЩЕМ	409
А. О. Mokrohuz MINIMAL HTTP HEADER CONFIGURATION FOR HTTP TRAFFIC OPTIMIZATION	411
О.М. Гайтан, А.С. Кікоть ДВОХФАКТОРНА АУТЕНТИФІКАЦІЯ КОРИСТУВАЧА КОМП'ЮТЕРНОЇ НАВЧАЛЬНОЇ ПРОГРАМИ ЗА ДАНИМИ КЛАВІАТУРНОГО ПОЧЕРКУ	413
О. Л. Ляхов, М. А. Мельник СТИСНЕННЯ ТЕКСТУ ЯК ЕТАП ПІДГОТОВКИ ДО АНАЛІЗУ	416
Ю.С.Ямненко, Є.О. Желязков ДИСКРЕТНІ СПЕКТРАЛЬНІ ПЕРЕТВОРЕННЯ УОЛША ТА АДАМАРА ТА ЇХ ЗАСТОСУВАННЯ ДЛЯ СИСТЕМ РОЗШИРЕННЯ СПЕКТРУ	420
О.Л. Ляхов, Н.А. Фурсова АЛГОРИТМ КЛАСТЕРИЗАЦІЇ ДЛЯ DATA MINING.....	425
А.О.Радченко ВИКОРИСТАННЯ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ РОЗПІЗНАВАННЯ АВТОМОБІЛЬНИХ НОМЕРІВ НА ПРАКТИЦІ.....	428
В.І. Гур'єв, І.В.Фірсова МОДЕЛЮВАННЯ ДВОЇСТОЇ ЗАДАЧІ ЗА ДОПОМОГОЮ Mathcad.....	432
М.О. Цирулін СИСТЕМА ЕЛЕКТРОННОЇ РЕЄСТРАЦІЇ ДЛЯ ОТРИМАННЯ МЕДИЧНИХ ПОСЛУГ «НІЖИН МЕДИЧНИЙ».....	436
Г.Д. Тулайдан ОСОБЛИВОСТІ ОПЕРАТИВНОГО УПРАВЛІННЯ ПІДПРИЄМСТВАМИ В УМОВАХ МЕРЕЖЕВИХ КІБЕРАТАК (ХАКЕРСЬКИХ АТАК).....	432

АЛГОРИТМ КЛАСТЕРИЗАЦІЇ ДЛЯ DATA MINING

О.Л. Ляхов, Н.А. Фурсова

*Полтавський національний технічний університет
імені Юрія Кондратюка*

В умовах загальної інформатизації задача кластеризації є однією із актуальніших задач Data Mining.

Є підстави вважати, що універсальний алгоритм кластеризації буде мати обчислювальну складність $O(2^n)$, де n – довжина вхідного слова. Проте складність задач цього типу іншого гатунку.

Задачі Data Mining полягають у знаходженні та візуалізації прихованої у даних інформації. Достатньо повна постановка задачі може бути здійснена тільки природною мовою з подальшою формалізацією на кроці алгоритмізації розв'язування. Враховуючи різноманіття областей, досвід та практику застосування методів Data Mining, можна стверджувати, що універсальний алгоритм кластеризації не існує.

На даний час розроблено багато алгоритмів кластеризації. Разом з цим, темпи поширення областей застосування Data Mining такі, що модифікація та розроблення нових алгоритмів кластеризації – актуальна проблема Data Mining (за даним пошукової системи Google, якщо кількість сторінок, присвячених цій тематиці, у 2005 році складала порядку 18 тисяч, то натеper - біля 60 тисяч, а у спорідненій галузі Big Date темпи зростання ще більші, порядку 280 тисяч сторінок, у чому легко упевнитися).

У даній роботі пропонується ідея алгоритму кластеризації, докладне дослідження цього алгоритму – предмет подальших публікацій.

Кластер – конструктивна відповідність такого поняття теорії множин, як «клас» або «властивість». Під "властивістю" у роботі розуміється деяке відношення феноменологічної симетрії на вихідній множині даних. Пропонується розглядати вихідну множину даних, як абстрактний комплекс. Тоді кластери – симплекси такого комплексу, а кластеризація полягає у означенні відношення, яке детермінує певну зв'язаність абстрактного комплексу, і у побудові алгоритму дослідження зв'язаності абстрактного комплексу, виходячи із заданого відношення.

Результат роботи алгоритму – ізольовані симплекси (кластери). У частинному випадку, коли кластер містить один елемент, це «викид», за загально прийнятою термінологією.

У даній роботі розглянутий простий приклад кластеризації множини елементів, кожній з яких поставлена у відповідність пара чисел, а кожній парі - евклідова відстань, яка є мірою їх схожості (близькості). Відомо, що кожному абстрактному комплексу над множиною з n елементів завжди можна поставити у відповідність геометричний комплекс у метричному просторі розмірності $n - 1$. При цьому задача кластеризації, так як вона поставлена, полягає у знаходженні симплексів, у яких кожний елемент схожий на решту в однаковій мірі.

Якщо множина, що спостерігається, складається з n ідентичних об'єктів, їй у просторі розмірності $n - 1$ ставимо у відповідність точку, а міру близькості елементів $p = 1$. Якщо ж елементи однаково не схожі між собою - правильний n -кутник з трикутними гранями, кожна вершина якого має ступень n , а міру близькості $p = 0$. Для $n = 4$ - це правильний тетраедр.

Відношення схожості визначимо так. За порогове значення схожості у даному прикладі береться середнє арифметичне довжин ребер геометричного комплексу

$$R_{av} = \frac{\sum_{i=1}^{C_n^2} R_i}{C_n^2}, \quad (1)$$

а кластер природно визначається, як симплекс, у якому $R_i < R_{av}$.

Тоді близькість елементів у такому кластеру з k вершинами

$$p = \frac{\sum_{i=1}^{C_k^2} (R_{av} - R_i)}{C_k^2 R_{av}}. \quad (2)$$

Таке відношення схожості задовольняє всім викладеним вище положенням.

Вихідні дані. Дани об'єкти кількістю $n=14$, кожному з яких поставлено у відповідність два числа:

[1,27,19], [2,11,46], [3,25,15], [4,36,27], [5,35,25], [6,10,43], [7,11,44],
[8,36,24], [9,30,10], [10,26,14], [11, 9,45], [12,33,23], [13, 27, 16], [14,10, 47].(3)

Перше число у кортежі – номер об'єкта у послідовності.

1-й крок. Обчислюємо матрицю взаємних відстаней.

2-й крок. Обчислюємо за формулою (1) $R_{av} = 20.41$.

3-й крок. Послідовно для всіх елементів матриці перевіряємо відношення $R_i < R_{av}$. Якщо «ні», викреслюємо стовпчик і рядок з тим ж номером.

4-й крок. Якщо викреслені є, утворюємо з них матрицю і повторюємо крок 3. Якщо ні, ідемо далі.

5-й крок. Для результатів кроку 3, за формулою (2) обчислюємо міру близькості. Якщо для якогось кластера вона порядку або менше порогового значення (у даному прикладі порогове значення прийнято

0.5), для цього кластера йти на крок 2. Якщо більша, то обчислення припиняються.

Результати роботи алгоритму у нашому прикладі наведені нижче у вигляді матриці взаємних відстаней, перший рядок якої – номери об'єктів згідно вихідним даним (3):

$$\begin{aligned}
 \text{Klas}_1 &:= \begin{bmatrix} 12 & 4 & 5 & 8 \\ 0. & 5. & 2.828 & 3.162 \\ 5. & 0. & 2.236 & 3. \\ 2.828 & 2.236 & 0. & 1.414 \\ 3.162 & 3. & 1.414 & 0. \end{bmatrix} & p1 := 0.86 \\
 \text{Klas}_2 &:= \begin{bmatrix} 1 & 3 & 10 & 13 \\ 0. & 4.472 & 5.099 & 3. \\ 4.472 & 0. & 1.414 & 2.236 \\ 5.099 & 1.414 & 0. & 2.236 \\ 3. & 2.236 & 2.236 & 0. \end{bmatrix} & p2 := 0.85 \\
 \text{Klas}_3 &:= \begin{bmatrix} 2 & 6 & 7 & 11 & 14 \\ 0. & 3.162 & 2. & 2.236 & 1.414 \\ 3.162 & 0. & 1.414 & 2.236 & 4. \\ 2. & 1.414 & 0. & 2.236 & 3.162 \\ 2.236 & 2.236 & 2.236 & 0. & 2.236 \\ 1.414 & 4. & 3.162 & 2.236 & 0. \end{bmatrix} & p3 := 0.88
 \end{aligned}$$

Об'єкт з номером 9 є викидом.

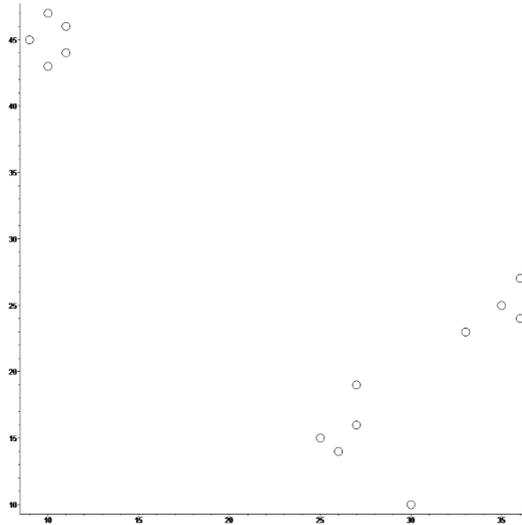


Рисунок. Візуалізація результатів роботи алгоритму кластеризації.

Висновки:

1. Запропоновано алгоритм, який за загально прийнятою класифікацією можна віднести до дивізимних алгоритмів.
2. Розглянуто простий приклад, що, з одного боку, дає змогу продемонструвати ідею алгоритму, а з іншого порівняти його з відомими, узагальненням якого його можна вважати.
3. У прикладі знаходяться кластери, утворені бінарними відношеннями $<$, $>$, $=$, а також, евклідова метрика. Він стає схожим на алгоритм BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) і т. і., але має суттєві відміни. Дозволяє, наприклад, заходити не тільки кластери, а й викиди.
4. У прикладі симплекси представляють собою повні графи. У загальному випадку симплекс - *k-арне* відношення з широким полем інтерпретацій. Алгоритм узагальнюється відповідним чином, але послідовність етапів зберігається.
5. У загальному випадку запропонований алгоритм полягає у дослідженні зв'язаності абстрактного комплексу, що надає змогу використовувати методи комбінаторної топології, виходячи із відношення близькості та зв'язаності, продиктованих конкретним класом задач. Докладне дослідження алгоритму, його переваги та недоліки, а також питання оптимізації – предмет подальших досліджень.

УДК 004.02

ВИКОРИСТАННЯ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ РОЗПІЗНАВАННЯ АВТОМОБІЛЬНИХ НОМЕРІВ НА ПРАКТИЦІ

А.О.Радченко

Чернігівський національний технологічний університет

Згорткова нейронна мережа – це штука нейронна мережа спеціальної архітектури, яка призначена для розпізнавання зображень. Вперше була запропонована в 1998 році Яном Лекуном (Yann LeCun)[1].

Для практичного застосування згорткових нейронних мереж в сторонньому ПО необхідні окремі бібліотеки, яких на сьогоднішній день безліч: `tiny-dnn`[2], `dyNet`[3], `mkl-dnn`[4], `convnet`[5], `pdnn`[6], `tensorflow`[7], `CNTK`[8] і т.д. Бібліотеки `tiny-dnn`, `dyNet`, `mkl-dnn`, `convNet` призначені для використання в програмах, написаних на C++, `pdnn` – на python, `tensorflow` і `CNTK` – на python і C++. При розробці прикладного ПО для розпізнавання автомобільних номерів однієї біблі-