

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧОРНОМОРСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ПЕТРА МОГИЛИ

КОЗАЧОК ЮРІЙ АНАТОЛІЙОВИЧ

УДК 004.925.5

**Використання зважених відстаней при застосуванні об'єктних кластерів
в нейронних мережах**

Спеціальність 123 – Комп'ютерна інженерія

Автореферат
магістерської роботи
на здобуття кваліфікації магістра з комп'ютерної інженерії

Миколаїв – 2019

Робота виконана у Чорноморському національному університеті ім. Петра Могили.

- Науковий керівник:** доктор технічних наук
Валентин Якович Кутковецький,
ЧНУ ім. Петра Могили,
професор кафедри комп'ютерної інженерії
- Рецензент:** Кандидат пед. наук
Болюбаш Надія Миколаївна,
ЧНУ ім. Петра Могили,
доцент кафедри ІС
- Консультант:** д-р біол. наук, професор
Томілін Юрій Андрійович,
ЧНУ ім. Петра Могили,
професор кафедри екології Медичного інституту

Захист відбудеться «27» лютого 2019 р. о 10⁰⁰ на засіданні Державної екзаменаційної комісії в ЧНУ ім. Петра Могили, ауд. 2-406

З магістерською роботою можна ознайомитись на сайті ЧНУ ім. Петра Могили за посиланням <http://chmnu.edu.ua>

Автореферат оприлюднений «25» лютого_2019 р.

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Кластерний аналіз насамперед виступає як важливий інструмент регіонального розвитку країни. Слід зауважити, що науковці ще недостатньо уваги приділяють застосуванню методів кластерного аналізу для вирішення конкретних практичних проблем, які кожного дня постають перед товаровиробниками. Результати таких досліджень знайшли своє відображення у працях Н.А. Волкової, А.В. Вороніна, С.С. Гаркавенко, А.Ю. Геворкяна, Г.М. Квіти, Н.В. Котелевської, А.С. Лавренка, Л.М. Попової, М.В. Рети, О.В. Стукача та ін. Кластерний аналіз є дуже важливою складовою у нейронній мережі для якості її роботи, тому у даній роботі запропоновано метод зважених відстаней між об'єктами, котрий працює краще у порівнянні з вище вказаними методами.

Мета та завдання дослідження. Розробити методику використання кластеризації у нейронних мережах з використанням зважених відстаней між об'єктами. Розробити власну нейронну мережу кластеризації, яка дозволяє підвищити точність роботи у порівнянні з існуючих систем кластеризації.

Для досягнення даної мети в магістерській роботі поставлені та вирішені наступні завдання:

Завдання:

- аналіз існуючих алгоритмів і методів кластеризації;
- аналіз ефективності зважених відстаней у вирішенні задач кластеризації;
- розробка методики кластеризації з використанням зважених відстаней;
- розробка нейронної мережі для вирішення задач кластеризації.

Об'єктом дослідження є процес впливу зваженої відстані на кластеризацію.

Предметом дослідження виступають методики кластеризації об'єктів зі зваженими відстанями між ними.

Методи дослідження: методи пошуку відстаней між об'єктами, методи кластеризації.

Структура та обсяг роботи. Магістерська робота складається з анотації на 2 сторінках, вступу, чотирьох розділів, висновків, переліку джерел посилання з 32 найменувань. Основна частина роботи становить 62 сторінки, серед яких 15 рис. та 19 табл.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** подано обґрунтування актуальності теми магістерської роботи, зазначено її зв'язок із науковою програмою, планами і темами, сформульовано мету та завдання дослідження, вказано практичне значення одержаних результатів, наведено відомості про апробацію результатів роботи та публікації автора. Зазначено, що спектр наукових досліджень на основі застосування методів кластерного аналізу дуже широкий: археологія, інформатика, медицина, психологія, хімія, біологія, державне управління, філологія, антропологія, маркетинг, соціологія тощо.

У **першому розділі** магістерської роботи «Огляд методів кластеризації» проведено огляд методик кластеризації, введено поняття зваженої відстані. Визначено, що в результаті застосування різних методів кластеризації можуть бути отримані неоднакові результати

Розглянуті основні класи задач, що вирішуються у межах цієї предметної сфери. У залежності від області застосування системи висувуються різні вимоги до точності і надійності (іноді за рахунок жертвуванням одним з цих показників заради покращення іншого). Озвучені завдання кластеризації та сформульовані задачі досліджень дипломної роботи.

Визначено, що кластеризація застосовується для того, щоб здійснити стиснення даних для скорочення обсягу використовуваних даних за рахунок того, що всередині кластера об'єкти не розрізняються (розглядаються як один об'єкт). Кластеризація може забезпечити краще розуміння даних і спростити їх подальшу обробку. Доведено, що проблема вимірювання близькості об'єктів постійно виникає при будь-яких тлумаченнях кластерів та різних методах класифікації

У другому розділі магістерської роботи «Технології розробки кластерного аналізу з використанням зважених відстаней» проведено аналіз методів реалізації, обрані технології для досягнення мети. Розроблена аналітична модель нейронної мережі з використанням зважених відстаней. Визначено, що оцінка зваженої відстані є дуже важливою у кластеризації, тому дану оцінку повинні робити експерти. Було запропоновано використовувати різні методи для оцінки, кожен з яких має свою ефективність (рис.1)

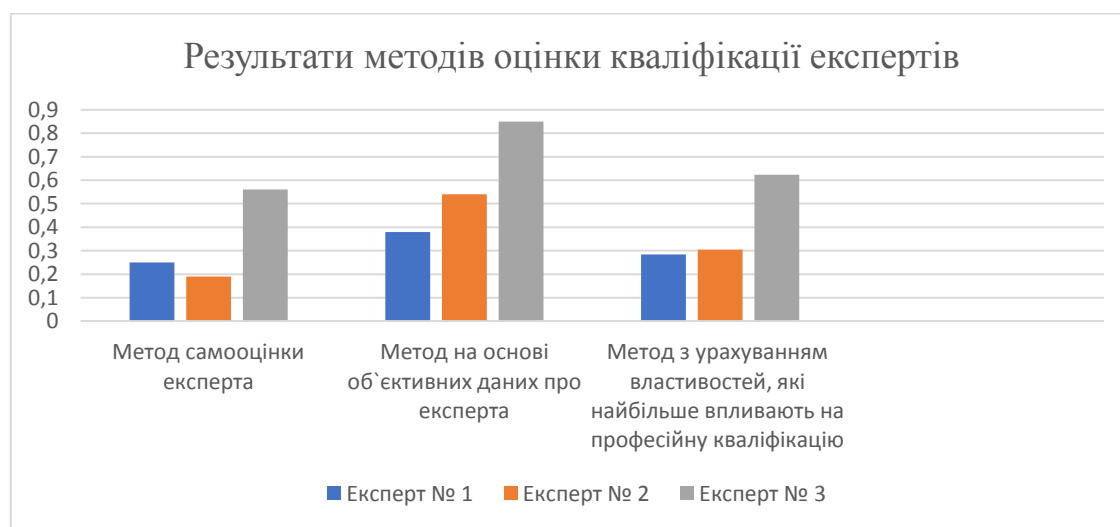


Рисунок 1

У третьому розділі магістерської роботи «Реалізація методики кластерного аналізу з використанням зважених відстаней» створено нейронну мережу та апробовано на двох задачах кластеризації:

- 1 Кластеризація банківських транзакцій, тобто виділення транзакцій що не є кіберзагрозою для банку та транзакцій що можуть нанести шкоду банківській безпеці;
- 2 Кластеризація студентів після складання іспиту за рівнем успішності. Виділено 4 кластери в залежності від ознак гендерної ознаки студентів та їх успішності.

Додатки містять лістинг коду ПЗ реалізації нейронної мережі зі зваженими відстанями.

Спеціальна частина «Охорона праці та безпека у надзвичайних ситуаціях» проведений аналіз факторів виробничого середовища у приміщенні на підприємстві ФОП «КОЗАЧОК Ю. А.», а також визначений вплив цих факторів на

здоров'я та працездатність працівників. Слід зазначити, що була встановлена відповідність всіх розглянутих показників чинним санітарним нормам та виявлено, що умови праці в ФОП « КОЗАЧОК А. О.» є оптимальними.

ВИСНОВКИ

У ході виконання магістерської роботи розроблено методики кластерного аналізу з використанням роширених і зважених відстаней між об'єктами та створено нейронну мережу, яка підвищила точність роботи у виконанні кластеризації та надала експерту, що проводить кластеризацію, більшої гнучкості у керуванні процесом розбиття об'єктів на кластери у порівнянні з існуючими системами.

Можна зробити висновок, що нейронні мережі для кластеризації можна задіювати в різних сферах дослідження: інформатика, економіка, медицина, безпека, тощо.

Внаслідок більш точного урахування впливу окремих змінних на функцію мети $F(x_1, x_2, \dots)$, отримані об'єктні кластери, які більш точно відображають вплив початкових змінних на досліджуваний процес.

Розроблена методика дозволяє скоротити матеріальні витрати і час у порівнянні з існуючими методами розбиття об'єктів по кластерам.

Алгоритм розробленої методики може бути використаний при створенні інших методик аналізу.

АНОТАЦІЯ

Магістерської наукової роботи на тему:

«Використання зважених відстаней при застосуванні об'єктних кластерів в нейронних мережах»

Студента 605 групи ЧНУ імені Петра Могили

Козачка Юрія Анатолійовича

Магістерська наукова робота присвячена розробленню методики кластерного аналізу з використанням зважених відстаней.

Метою роботи є розроблення методики використання кластеризації у нейронних мережах з використанням зважених відстаней між об'єктами та створення власної нейронної мережі кластеризації, яка дозволяє підвищити точність роботи у порівнянні з існуючими системами кластеризації.

Об'єкт дослідження: процес впливу зваженої відстані на кластеризацію.

Предмет дослідження: методики кластеризації об'єктів зі зваженими відстанями між ними.

Для розв'язання поставлених завдань використано наступні методи наукового дослідження: теоретичний аналіз наукових літературних джерел, огляд існуючих аналогів. Пояснювальна записка дипломної роботи складається зі вступу, трьох розділів, висновків, методичної частини та спеціальної частини з охорони праці. У вступі визначається актуальність теми та проводиться короткий огляд поставленої задачі.

В першому розділі виконується аналіз існуючих методик кластеризації. У другому розділі проводиться вибір програмного забезпечення для вирішення поставленої мети. У третьому розділі наведено опис розробленої методики та результати її тестування. У висновках проводиться аналіз проведеної роботи та отриманих результатів.

У спеціальній частині з охорони праці йдеться про техніку безпеки на робочому місці працівника кафедри інтелектуальних інформаційних систем ЧНУ імені Петра Могили.

Дипломна робота містить 82 сторінки (без додатків), 15 рисунків, 18 таблиць.

Ключові слова: нейронні мережі, кластерний аналіз, зважені вершини, нейронна мережа Коханена, компетентність експертів.

ABSTRACT

of master's scientific work on topic:

"The use of weighted distances in the application of object clusters in neural networks"

Student of group No. 605 of Petro Mohyla Black Sea National University

Yurii Kozachok

Master's scientific work is devoted to the development of methods of cluster analysis using weighted distances.

The aim of the work is to develop a method for using clustering in neural networks using weighted distances between objects and their own neural network of clustering which allows to improve the accuracy of work compared with existing clustering systems.

Object of research: process of influence of weighted distance on clusterization.

Subject of research: methods of clustering objects with weighted distances between them.

The following methods of scientific research were used to solve the tasks: theoretical analysis of scientific literary sources, review of existing analogues.

The explanatory note of the thesis consists of an introduction, three sections, conclusions, a methodological part and a special part on labor protection.

The introduction determines the relevance of the topic and provides a brief overview of the task.

The first section analyzes the existing clustering techniques. In the second section, a choice of software is being made to solve the set goal. The third section describes the developed methodology and the results of its testing. In the conclusions, the analysis of the work performed and the results obtained.

The conclusions consist of the analysis the work and the results.

The project's specialized part considers a occupational safety refers to safety in the workplace specialists of employees of computer department.

This graduation project contains 82 pages (without attachments), 15 figures, 18 tables.

Keywords: neural networks, cluster analysis, weighted vertices, neural network, beloved, expert competence.