

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ІВАНА ПУЛЮЯ  
ФАКУЛЬТЕТ КОМП'ЮТЕРНО-ІНФОРМАЦІЙНИХ СИСТЕМ  
І ПРОГРАМНОЇ ІНЖЕНЕРІЇ  
КАФЕДРА КОМП'ЮТЕРНИХ НАУК

**ДОБРУЦЬКИЙ КОСТЯНТИН ВІКТОРОВИЧ**

УДК 004.04

МЕТОДИКА ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ПРОЦЕСУ ПРИЙНЯТТЯ РІШЕНЬ  
ЗА РАХУНОК ВПРОВАДЖЕННЯ ГІБРИДНОЇ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ  
ДЛЯ РІШЕННЯ ЗАДАЧ ВИДОБУВАННЯ ДАНИХ

122 «Комп'ютерні науки та інформаційні технології»

**Автореферат**

дипломної роботи на здобуття освітнього ступеня «магістр»

Тернопіль  
2018

Роботу виконано на кафедрі комп'ютерних наук Тернопільського національного технічного університету імені Івана Пулюя Міністерства освіти і науки України

**Керівник роботи:** кандидат економічних наук, доцент кафедри комп'ютерних наук  
**Матійчук Любомир Павлович,**  
Тернопільський національний технічний університет імені Івана Пулюя,

**Рецензент:** кандидат технічних наук, доцент кафедри комп'ютерно-інтегрованих технологій  
**Карташов Віталій Вікторович,**  
Тернопільський національний технічний університет імені Івана Пулюя

Захист відбудеться \_19\_ лютого 2018 р. о 9-00 годині на засіданні екзаменаційної комісії №30 у Тернопільському національному технічному університеті імені Івана Пулюя за адресою: 46001, м. Тернопіль, вул. Руська 56, навчальний корпус №1, ауд. 702

## ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

**Актуальність теми роботи** дана робота присвячена розробці методики для побудови гібридної інтелектуальної системи для рішення основного класу задач Data Mining та дослідженню методів і алгоритмів кластеризації та регресії. Інтелектуальний аналіз даних пов'язаний з пошуком в даних прихованих нетривіальних і корисних закономірностей, що дозволяють отримати нові знання про досліджувані дані. Особливий інтерес до методів аналізу даних виник у зв'язку з розвитком засобів збору і зберігання даних. У фахівців з різних галузей людської діяльності постає питання про обробку даних. Статистичні методи покривають лише частину потреб з обробки даних. У такій ситуації методи інтелектуального аналізу даних набувають особливої актуальності. Їх головна особливість полягає у встановленні наявності та характеру прихованих закономірностей у даних, тоді як традиційні методи займаються головним чином параметричною оцінкою вже встановлених закономірностей. Серед методів інтелектуального аналізу даних особливе місце займають класифікація, кластеризація та регресія. Задача кластеризації має різні способи вирішення. Складність полягає у відсутності на момент початку аналізу будь-якої додаткової інформації про дані. Тому особливу актуальність має розроблення методики адаптивної кластеризації, за якої вибір найкращого рішення здійснюється формально, за заданими критеріями.

**Мета роботи:** підвищення ефективності процесу прийняття рішень за рахунок впровадження гібридної інтелектуальної системи для рішення основного класу задач Data Mining.

**Об'єкт, методи та джерела дослідження.** Процес побудови гібридної інтелектуальної системи для вирішення основного класу задач Data Mining.

**Наукова новизна отриманих результатів:** запропонована методика побудови гібридної інтелектуальної системи для рішення основних задач Data Mining за допомогою кластеризації та регресії.

**Практичне значення отриманих результатів.** Розробка методики адаптивної кластеризації та регресії, заснованої на використанні критеріїв оцінки якості рішення і яка дозволить повністю формалізувати рішення задачі кластеризації та регресії. Відмінною особливістю методики є оцінка якості кожного розбиття та вибір найкращого з них.

**Апробація.** Окремі результати роботи доповідались на XX науковій конференції Тернопільського національного технічного університету імені Івана Пулюя, Тернопіль, ТНТУ, 17 – 18 травня 2017 р., та VI міжнародна науково-технічна конференція молодих вчених та студентів науковій конференції Тернопільського національного технічного університету імені Івана Пулюя, Тернопіль, ТНТУ, 16 – 17 листопада 2017 р.

**Структура роботи.** Робота складається з розрахунково-пояснювальної записки та графічної частини. Розрахунково-пояснювальна записка складається з вступу, 7 частин, висновків, переліку посилань та додатків. Обсяг роботи: розрахунково-пояснювальна записка – 148 арк. формату А4, графічна частина – 9 аркушів формату А1

## ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі проведено огляд основних проблем інтелектуальному аналізу даних та проведено теоретичний та практичний аналіз вихідних необроблених даних в цінну інформацію для прийняття стратегічних рішень.

В розділі «Аналіз гібридних інтелектуальних систем та проблеми їх розробки» розглянуто комплексність проблем кластеризації та регресії, яка диктує необхідність міждисциплінарного підходу до їх вирішення, розширення та інтеграцію вже наявних методик. Принциповою відмінністю парадигми проектування гібридних систем від традиційної побудови інтелектуальних обчислювальних систем є можливість синтезувати якісно різні варіанти рішень. Даний синтез в рівній мірі залежить як від специфіки поставленої задачі, так і від внутрішніх параметрів системи, умов зовнішнього середовища, а також характеру їх взаємодії. Це дає підставу розглядати застосування гібридних інтелектуальних систем як стратегічний напрям у галузі інтелектуального аналізу даних..

В розділі «Розроблення методики побудови гібридної інтелектуальної системи для вирішення основного класу задач Data Mining» Розроблена методика побудови гібридної інтелектуальної системи для рішення основних задач Data Mining, а саме кластеризації та регресії веде до підвищення ефективності аналізу статистичної інформації. Тому запропонована методика може знайти застосування у багатьох галузях.

Формалізований метод побудови гібридної інтелектуальної системи для рішення основних задач Data Mining, а саме кластеризації та регресії дозволяє кластери зі схожими об'єктами, а потім відносно побудованих кластерів проводити регресійний аналіз, тобто проводити аналіз взаємозв'язків.

В розділі «Проектування гібридної інтелектуальної системи для рішення задач кластеризації та регресії» розраховані показники дозволили проаналізувати якість та ефективність побудови гібридної інтелектуальної системи. Вирішувати задачу регресії значно ефективніше, якщо попередньо вибірку згрупувати за однорідністю даних, та провести регресійний аналіз для кожного кластера окремо, ніж для всіх даних в сукупності..

В спеціальній частині розглянуто питання: моделі Data Mining, статистичні методи Data Mining та кібернетичні методи Data Mining..

В частині «Обґрунтування економічної ефективності» розглянуто питання організації виробництва і проведено розрахунки техніко-економічної ефективності проектних рішень.

В частині «Охорона праці та безпека в надзвичайних ситуаціях» розглянуті питання по законодавчій базі охорони праці в галузі інформаційних технологій, а також вимоги з охорони праці при роботі з ПК. В розділі безпека в надзвичайних ситуацій розглянуто питання: безпеки в соціально-економічній сфері, заходи щодо зменшення негативного впливу персональних комп'ютерів та землетруси і порядок дій населення при них.

В частині «Екологія» розглянуто етапи та техніка збору та опрацювання екологічної інформації та методи узагальнення екологічної інформації

У загальних висновках щодо дипломної роботи наведено отримані технічні рішення і запропоновано організаційно-технічні заходи, які забезпечують виконання поставленого завдання.

## ВИСНОВКИ

Виконаний огляд літературних джерел показав, що існує багато підходів до побудови гібридних інтелектуальних систем. Існуючі системи підтримки прийняття рішень використовуються у сфері управління бізнесом, у промисловості, на виробництві, одним словом там, де існує необхідність прийняття рішень, особливо коли йде мова про неструктуровані і слабкоструктуровані завдання.

Комплексність проблем кластеризації та регресії диктує необхідність міждисциплінарного підходу до їх вирішення, розширення та інтеграцію вже наявних методик. Принциповою відмінністю парадигми проектування гібридних систем від традиційної побудови інтелектуальних обчислювальних систем є можливість синтезувати якісно різні варіанти рішень. Даний синтез в рівній мірі залежить як від специфіки поставленої задачі, так і від внутрішніх параметрів системи, умов зовнішнього середовища, а також характеру їх взаємодії. Це дає підставу розглядати застосування гібридних інтелектуальних систем як стратегічний напрям у галузі інтелектуального аналізу даних.

Спроектвана система зможе широко застосовуватися підтримки прийняття рішень. Для побудови такої системи запропоновано використовувати дворівневу архітектуру. На верхньому рівні використовуючи агломеративні та ітеративні методи кластерного аналізу групуючи вхідні дані в кластери. На нижньому рівні проводити регресійний аналіз як для всієї вибірки так і для кожного кластеру окремо. Це доцільно для підвищення надійності, масштабованості та інтегрованості.

Сформульована за допомогою математичного апарату гібридна інтелектуальна система дозволяють побудувати гібридну інтелектуальну систему для підвищення ефективності прийняття рішень.

Виконаний аналіз адекватності запропонованої методики побудови гібридної інтелектуальної системи показав відповідність реальним даним. Проведений аналіз ефективності по зрівнянню із існуючими аналогами показав, що похибка моделі, побудованої з використання ієрархічного підходу та об'єднанням відповідних методів, дозволяє отримати більш ефективні результати, ніж за допомогою існуючих методів. Таким чином, проведені експерименти довели можливість його практичного застосування в різних сферах.

Отже, впровадження гібридної інтелектуальної системи для підтримки прийняття рішень може підвищити ефективність прийняття рішень.

## СПИСОК ОПУБЛІКОВАНИХ АВТОРОМ ПРАЦЬ ЗА ТЕМОЮ РОБОТИ

1. Добруцький К. Огляд та порівняльний аналіз алгоритмів та методів кластеризації та регресії КС [Текст] / К. Добруцький, О. Назаревич. Тези доповіді на XX наукова конференція Тернопільського національного технічного університету імені Івана Пулюя – Тернопіль, ТНТУ, 2017. – с. 75.

2. Добруцький К. Проблеми захисту комп'ютерних систем. [Текст] / К. Добруцький, Г. Шимчук., В. Юрків Тези доповіді на VI міжнародну науково-технічну конференцію молодих учених та студентів Тернопільського національного технічного університету імені Івана Пулюя – Тернопіль, ТНТУ, 2017. – с. 56-57.

### АНОТАЦІЯ

Пояснювальна записка дипломної роботи ОКР «Магістр» містить результати розробки гібридної інтелектуальної системи для рішення задач кластеризації та регресії. Проведено аналіз існуючих систем, розглянуто їх класифікацію та архітектуру. Було досліджено підходи до побудови гібридних інтелектуальних систем. Розроблені методи побудови гібридної інтелектуальної системи.

Наведені в роботі критерії дозволяють оцінити якість рішень задачі кластеризації та регресії, одержуваних як за допомогою традиційних методів, так і за допомогою нового методу запропонованого в роботі; описана в роботі методика адаптивної кластеризації та регресії та рекомендації щодо її застосування дозволяють значно формалізувати аналіз даних.

**Ключові слова:** ГІБРИДНА ІНТЕЛЕКТУАЛЬНА СИСТЕМА, КЛАСТЕРИЗАЦІЯ, РЕГРЕСІЯ, ІНТЕЛЕКТУАЛЬНА ОБРОБКА ІНФОРМАЦІЇ, DATA MINING, ПІДТРИМКА ПРИЙНЯТТЯ РІШЕНЬ, МОДЕЛЬ, ГІБРИД, БАГАТОРІВНЕВИЙ ПІДХІД, ІНФОРМАЦІЯ

### ANNOTATION

Explanatory note of the dissertation "Master" contains the results of the development of a hybrid intelligence system for solving the problems of clustering and regression. The analysis of existing systems, their classification and architecture is considered. Approaches to the construction of hybrid intelligence systems were explored. Methods of constructing a hybrid intelligence system are developed.

The criteria given in this work allow us to estimate the quality of the solutions of the clusterization and regression problem obtained both by traditional methods and by the new method proposed in the work; the methodology of adaptive clustering and regression described in the work and recommendations for its application make it possible to significantly formalize data analysis.

**Key words:** HYBRID INTELLEAGENT SYSTEM, CLUSTERING, REGRESSION, INTELLEAGENT INFORMATION PROCESSING, DATA MINING, DECISION SUPPORT SYSTEM, DECISION MAKING, SUPPORT, INFORMATION, HYBRID, MULTILEVEL APPROACH