

**Замрій І.В.***Державний університет телекомунікацій, Київ*

## ТЕХНОЛОГІЇ МАШИННОГО НАВЧАННЯ ФУНКЦІОНАЛЬНО СТІЙКОЇ ІНТЕЛЕКТУАЛЬНОЇ ІНФОРМАЦІЙНОЇ СИСТЕМИ ЗАКЛАДУ ВИЩОЇ ОСВІТИ

**Анотація:** Розвиток ринку породило потребу в широко функціональних інтегрованих корпоративних інформаційних системах, в яких поєднуються бази для зберігання даних, аналітичні інструменти, і системи документообігу. В сучасних умовах питання щодо вибору корпоративної інформаційної системи часто є ключовим стратегічним рішенням, багато в чому визначає ефективність діяльності підприємства. Тому розробка, підтримка та оптимізація інтелектуальних інформаційних систем є актуальним питанням забезпечення оптимального прийняття рішень на основі аналізу поточних ситуацій для досягнення певної мети. З цією метою розглянуто оптимізацію алгоритмів кластеризації інтелектуального аналізу даних на основі хмарних технологій і машинного навчання для підвищення ефективності інтелектуальної інформаційної системи закладу вищої освіти.

Вибір найкращого алгоритму за умов конкретного завдання має обґрунтовано здійснюватися особою, яка приймає рішення, оскільки вкрай важливим є процеси інтерпретації та оцінювання здобутих результатів після їх аналізу. На даній стадії головну роль відіграє експерт у досліджуваній предметній галузі, який крім використання критеріїв може на основі апріорних уявлень і знань ключових цільових показників здійснити додаткову верифікацію результатів для подальшого прийняття рішень. Для досягнення максимального результату потрібен комплексний підхід до аналізу даних, що включає в себе як використання апріорних знань фахівців для попереднього оброблення даних і інтерпретації результатів, так і застосування спеціалізованих алгоритмів.

У роботі проведено аналіз недоліків, виконано оптимізацію та розпаралелювання оптимізованих алгоритмів, щоб покращити здатність обробляти великий масив даних та підвищити ефект від діяльності інтелектуальної інформаційної системи закладу вищої освіти за допомогою використання хмарних обчислень та машинного навчання інтелектуальної інформаційної системи.

**Ключові слова:** інтелектуальна інформаційна система, функціональна стійкість, алгоритми кластеризації інтелектуального аналізу даних, машинне навчання, заклад вищої освіти.

**Zamrii I.V.***State University of Telecommunications, Kyiv*

## MACHINE LEARNING TECHNOLOGIES OF THE FUNCTIONALLY SUSTAINABLE INTELLECTUAL INFORMATION SYSTEM OF THE INSTITUTION OF HIGHER EDUCATION

**Abstract:** Market development has created a need for widely functional integrated corporate information systems that combine data storage databases, analytical tools, and document management systems. In today's conditions, the question of choosing a corporate information system is often a key strategic decision that largely determines the efficiency of the enterprise. Therefore, the development, support and optimization of intelligent information systems is an urgent issue of

*ensuring optimal decision-making based on the analysis of current situations to achieve a certain goal. For this purpose, the optimization of clustering algorithms of intelligent data analysis based on cloud technologies and machine learning to increase the efficiency of the intelligent information system of a higher education institution is considered.*

*The choice of the best algorithm under the conditions of a specific task should be reasonably carried out by the person who makes the decision, since the processes of interpretation and evaluation of the obtained results after their analysis are extremely important. At this stage, the main role is played by an expert in the subject field under investigation, who, in addition to using the criteria, can, based on a priori ideas and knowledge of key target indicators, perform additional verification of the results for further decision-making. To achieve the maximum result, a complex approach to data analysis is required, which includes both the use of a priori knowledge of specialists for pre-processing of data and interpretation of results, and the use of specialized algorithms.*

*The paper analyzed the shortcomings, optimized and parallelized the optimized algorithms in order to improve the ability to process a large array of data and increase the effect of the activity of the intelligent information system of the institution of higher education using cloud computing and machine learning of the intelligent information system.*

**Keywords:** *intelligent information system, functional stability, clustering algorithms of intelligent data analysis, machine learning, higher education institution.*

## 1. Вступ.

У зв'язку із зростанням обсягу ринку системної інтеграції та трансформації щодня з'являється все більше пропозицій з використання ІТ-інфраструктури для ефективного управління не лише підприємством, а й навіть закладу вищої освіти. Широкого використання набули розподілені корпоративні інформаційні системи, організація роботи яких залежить від безперебійної роботи створених на місцях центрів обробки інформації. Центри призначені для автоматизації процесів збирання та пересилання інформації, її аналізу та створення звітів з використанням економіко-математичних методів, технічних засобів та організаційної структури, яка забезпечує раціональне керування складними об'єктами й інформаційними процесами. З використанням інформаційних систем можна розв'язувати задачі первинного та оперативного планування виробництва, оперативного розподілу завантаження обладнання та розподіленого використання ресурсів організації.

Під функціональною стійкістю об'єкту [1-2] розуміється його властивість зберігати упродовж заданого часу виконання своїх основних функцій в межах, встановлених нормативними вимогами, в умовах впливу потоків відмов, несправностей, збоїв. Дана властивість тісно пов'язана з властивостями стійкості, надійності, живучості і відмовостійкості.

Стратифіковані моделі складних систем використовуються з метою моделювання і грають фундаментальну роль для правильного розуміння складної системи. Вони дозволяють більш наглядно зрозуміти сенс і призначення усієї системи через призначення підсистем. Звертаючись до нижніх страт, можна точніше і детально пояснити, яким чином система функціонує, як здійснюється та або інша конкретна операція. При русі вгору за ієрархією опис стає ширшим, охоплюючи більше число підсистем і великі періоди часу. Інтелектуальна інформаційна система управління в цілому може розглядатися як ієрархічно організована сукупність, яка складається з елементів технічного обладнання та програмного забезпечення. В такому випадку, архітектура інтелектуальних інформаційних систем управління закладу вищої освіти припускає багаторівневу організацію, де кожен з рівнів може бути представлений як певна абстрактна система, що виконує свої певні функції. Така структура поділяє складний механізм зв'язку і спрощує архітектуру до функціональних рівнів, що легко розуміються, а також дозволяє обмінюватися інформацією між системами, що мають різні характеристики і пов'язаними з різними типами систем. Кожен рівень, окрім самого верхнього, забезпечує обслуговування більш високого рівня, користуючись послугами нижчого рівня. Кожен з рівнів

однієї системи може бути пов'язаний з відповідним рівнем іншої системи за допомогою відповідних угод і протоколів. Інтерфейс між рівнями не обов'язково може бути жорстким, що дозволяє гнучко налаштовувати систему передачі інформації під різні завдання реального часу.

При ієрархічному підході [3] організації засобів забезпечення функціональної стійкості автоматизованої системи управління закладом вищої освіти проблема розподілу завдань, які повинні виконуватися різними рівнями або окремими елементами, вирішується на основі системного підходу до засобів забезпечення функціональної стійкості в цілому і до завдання, які ці засоби повинні виконувати. В цьому випадку використовуються концепції страти і шару, тобто з одного боку відбувається стратифікація моделі засобів забезпечення функціональної стійкості бездротової мережі, а з іншої – здійснюється декомпозиція, що ставиться перед засобами забезпечення функціональної стійкості, розбиття на шари. Завдання елементам, що утворюють багаторівневу систему, в цьому випадку визначають по відношенню до моделей і вирішуваних проблем, що з'являються на відповідній страті або шарі. При цьому завдання для декількох ешелонів можуть бути визначені з моделі однієї і тієї ж страти, а вирішувана проблема на цьому шарі може бути розподілена між рядом ешелонів, внаслідок чого відбувається вкладення одних ієрархій в інші. В такому випадку істотно спрощується завдання аналізу функціонально стійких інтелектуальних інформаційних систем підприємства.

## **2. Аналіз літературних даних і постановка проблеми**

Хмарні обчислення створили незліченну кількість нових способів роботи та бізнес-моделей. Спільне застосування Інтернет-технологій і хмарних обчислень сприяє комплексному збору та інтелектуальній обробці даних про час і покращує рівень управління інформацією в цій галузі [4].

Такі системи як розумний офіс — це нова модель офісу, яка використовує технологію хмарних обчислень для інтелектуального керування програмним і апаратним обладнанням, необхідним для бізнесу, і реалізує уніфіковане розгортання та доставку корпоративного прикладного програмного забезпечення [5]. Розумний офіс підтримує безпечний віддалений доступ кількох термінальних пристроїв, таких як персональні комп'ютери, мобільні телефони та планшетні комп'ютери, що може підвищити безпеку, простоту використання та масштабованість середовища корпоративного офісу, покращити співпрацю та спільне використання ресурсів, а також всебічно покращити ефективність управління та оптимізації бізнес-процесів, зниження операційних витрат [6]. Розумний офіс має широкий спектр простору для розробки та перспектив застосування в майбутньому. Системи розумного офісу компенсують більшість недоліків [7] та мають переваги у зниженні витрат на управління підприємством, підвищенні рівня та ефективності управління тощо [8-9].

Ефективне користування технологією хмарних обчислень та машинного навчання має низку позитивних моментів. Серед них можна назвати: високий рівень відмовостійкості, гнучкість конфігурації під потреби певного проекту, економія часу на організацію інфраструктури, зручність та простота у керуванні серверними ресурсами в єдиному інтерфейсі, зміна та докомплектація потужностей без зупинки “хмари” та автоматизація всіх ІТ-процесів. Відомо, що процес автоматизації розпочинається з виділення ресурсної потужності, підтримки відмовостійкості, модернізації платформи до розвитку віртуального дата-центру клієнтів.

Успішні хмарні рішення слід застосовувати як зміну або розширення традиційних програмних рішень для технічної та фундаментальної реорганізації бізнес-моделі. За їх допомогою можна зберегти надвеликі обсяги даних, доступ до яких здійснюється з мобільних та персональних комп'ютерів.

Інтелектуальна інформаційна система закладу вищої освіти — це комп'ютерна програма, яка забезпечує збір, аналіз та обробку даних про студентів, викладачів та інші аспекти управління університетом. Ця система включає в себе базу даних про студентів, їх успішність, розклад занять, науково-педагогічних працівників та їх кваліфікацію, наукові дослідження та

інші аспекти життя закладу вищої освіти. Основними компонентами інтелектуальної інформаційної системи закладу вищої освіти є база даних, алгоритми аналізу даних та інтерфейс користувача. Система може використовувати різні методи аналізу даних, зокрема актуальні такі як статистичний аналіз, машинне навчання та штучний інтелект. Інтелектуальна інформаційна система закладу вищої освіти може бути використана для розв'язання різних завдань, наприклад, прогнозування кількості студентів на наступний навчальний рік, виявлення проблем з успішністю студентів, аналізу ефективності науково-педагогічних працівників та інших показників. Основною перевагою інтелектуальної інформаційної системи є те, що вона забезпечує швидкий та точний аналіз даних, що дозволяє закладу вищої освіти приймати кращі рішення та дозволяє допомогти зменшити витрати часу і ресурсів для виконання завдань.

У цілому, інтелектуальна інформаційна система закладу вищої освіти є потужним інструментом для управління та прийняття рішень університетом. Вона дозволяє збільшити ефективність та точність процесів, що може позитивно вплинути на якість навчання та дослідження закладу вищої освіти.

Метою статті є аналіз інформаційних технологій та алгоритмів аналізу масивної інформації даних, робота яких дає змогу підвищити ефект від діяльності інформаційних систем, обґрунтування доцільності використання хмарних обчислень та машинного навчання інтелектуальної інформаційної [10], зокрема і для забезпечення функціональної стійкості інтелектуальної інформаційної системи закладу вищої освіти [11-13].

### **3. Дослідження алгоритмів аналізу масивної інформації даних**

Оскільки проблема поглибленого аналізу масивної інформації даних стала актуальною проблемою, яку необхідно вирішити, алгоритм кластеризації на основі хмарних обчислень розроблено та застосовано до масивної обробки даних інтелектуального офісу для підвищення ефективності роботи інтелектуальної інформаційної системи. Кластеризація означає повторне використання максимально можливої кількості ручної ідентифікаційної роботи, виконаної попередниками, для підвищення ефективності роботи. Вирішити формальну проблему кластеризації інтелектуального аналізу даних нескладно. Найпростішим і прямим способом є розробка конвертера між різними основними форматами даних. Оскільки популярних форматів даних небагато, а правила перетворення зрозумілі, це проблема, яку можна вирішити, якщо інвестувати певну кількість робочої сили. Однак кластеризація семантичного інтелектуального аналізу даних є більш складною, тому ця стаття проводить дослідження алгоритмів кластеризації інтелектуального аналізу даних на основі хмарних обчислень.

Кластерний аналіз полягає в класифікації наборів даних на основі принципів подібності. Кластеризація — це алгоритм інтелектуального аналізу даних, і це центральна галузь інтелектуального аналізу даних. Дослідники завжди вважали алгоритми кластеризації важливим напрямком досліджень, і алгоритми кластеризації застосовувалися в багатьох галузях великих компаній. Крім того, він відіграє важливу роль у групуванні тексту, класифікації груп користувачів, даних та пошуку. Алгоритм кластеризації може ефективно відкривати внутрішні закони речей, тому це важливий крок у розробці даних. Крім того, алгоритм кластеризації може розділити набір даних на кілька класів за певними правилами та вибрати відповідну формулу вимірювання подібності; таким чином, дані, що належать до одного класу, подібні один до одного, але дані між класами не схожі.

Потік алгоритму кластеризації показаний на рис. 1.

Існують різні методи кластеризації, такі як ієрархічна кластеризація, k-середніх, DBSCAN та інші. Кожен з цих методів має свої переваги та недоліки, тому вибір методу кластеризації залежить від конкретного завдання та характеристик даних. Перевагами кластеризації є можливість виявлення структури та закономірностей в даних, що дозволяє зробити більш точні прогнози та приймати кращі рішення. Крім того, кластеризація може бути використана для зменшення обсягу даних, що сприяє покращенню ефективності обробки

даних. Однак, кластеризація має свої недоліки, такі як складність вибору оптимального методу та параметрів кластеризації, можливість виникнення перекривання між кластерами та інші. Також, результати кластеризації можуть бути недостатньо точними, якщо дані мають велику кількість шуму або випадкових значень. Але у цілому, кластеризація є потужним інструментом для аналізу та обробки даних, що може бути використаний у різних областях, таких як маркетинг, медицина, наука, освіта та інші. Вона дозволяє виявляти структуру та закономірності в даних, що може позитивно вплинути на прийняття рішень та покращення ефективності процесів.



Рис.1. Процес кластеризації

У машинному навчанні алгоритм кластеризації є алгоритмом неконтрольованого навчання, тобто немає необхідності знати тип даних перед кластеризацією. Метою кластеризації є класифікація подібних об'єктів у наборі даних до однієї категорії. Загалом, оцінка ефекту кластеризації визначається відповідним алгоритмом кластеризації, приблизною відстанню вибірки та функцією оцінки. Коли кількість ітерацій досягає певного значення і функція судження прагне до нескінченності, кластеризація досягає найкращого ефекту [6].

Розглянемо два алгоритми.

Алгоритм 1:

Введення:  $k$  значення та набір даних  $D (D_1, D_2, \dots, D_i, \dots, D_n)$ .

Вихід: результати кластеризації.

- 1) Алгоритм вибирає випадковим чином  $k$  дані в наборі даних  $D$  як початкову центральну точку кластера.
- 2) Цей алгоритм знаходить відстань від усіх решти точок даних до  $k$  центру кластера і вибирає найближчу до себе точку як центр кластера.
- 3) Алгоритм обчислює центральні точки всіх класів по черзі та використовує ці центральні точки як нові центри кластерів.
- 4) Алгоритм повторює кроки 2) і 3) неодноразово, доки помилка не буде середня і значення  $E = \sum_{i=1}^k \sum_x |x - \bar{x}_i|^2$  менше вказаного.

Алгоритм 2

Вхід: перехоплення  $d_c$ .

Вихід: результати кластеризації.

- 1) Алгоритм знаходить відстань між усіма даними.
- 2) Алгоритм обчислює локальну щільність  $\rho_i$  усіх точок даних.
- 3) Алгоритм обчислює мінімальну відстань високої щільності  $\delta_i$  усіх даних.
- 4) Алгоритм знаходить центральну точку кластера на графі рішень.
- 5) Алгоритм групує решту точок до найближчої центральної точки.
- 6) Алгоритм видаляє точки шуму.

Ідея алгоритму 1 —  $K$ -середніх полягає в наступному: на першому кроці  $k$  значень випадковим чином вибираються набір даних як центр кластера; на другому кроці — відстань від кожного елемента до центральної точки обчислюється, і ці об'єкти вибирають найближчу до себе точку як центр класу; на третьому кроці середнє значення всіх об'єктів загалом класів обчислюється по порядку, а середнє значення використовується як центр нового класу; на четвертому кроці середнє помилка та значення  $E$  обчислюються, доки воно не стане меншим за заданий поріг; інакше виконується другий крок і ітерація продовжується.

Випадковий вибір  $k$  точок даних як початковий центр точки кластеризації зроблять ефект кластеризації нестабільним. Для наборів даних, які не є опуклими, випадковість початкових центрів зазвичай викликають локальне оптимальне рішення. Тоді ця функція має багато локальних оптимальних розв'язків, і тільки одне з багатьох локальних оптимальних рішень є глобальним оптимальним рішенням. Як показано на рис.2, оскільки початкові центральні точки кластеризації різні, графік цільової функції кластеризації буде слідувати чотирма різними шляхами  $V_a, V_b, V_c, V_d$  для досягнення локального оптимального рішення. Серед чотирьох локальних оптимальних рішень, тільки одне є глобальним оптимальним рішенням  $C$  отримане шляхом ітерації  $V_c$ -шляху. Таким чином, якщо вихідний центр кластеризації не вибирається відповідно, алгоритм кластеризації  $K$ -середніх зазвичай буде не знаходити глобальної оптимальної точки і залишатися на локальній оптимальній точці.

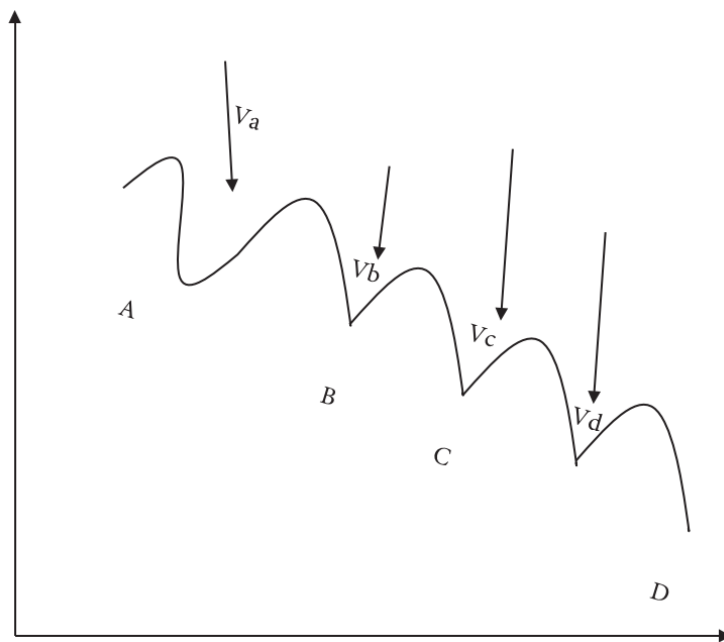


Рис.2. Просторовий стан функції критерію суми квадратів помилок.

Крім того, необхідно вказати кількість кластерів перед кластеризацією за алгоритмом 1. Однак буде багато наборів даних, які мають невідомі характеристики. Якщо кількість класів наборів даних не можна чітко вказати, то необхідно виконати кілька тестів кластеризації з різною кількістю кластерів  $k$ , тоді  $k$  вибирають з найкращим ефектом кластеризації за кількістю кластерів. Це число не випадкове і експерименти показують, що воно належить відрізку  $[2, \text{int}\sqrt{n}]$  для отримання оптимального рішення для функції ефекту кластеризації. Однак алгоритм  $K$ -середніх буде запущено кілька разів у процесі пошуку найкращого

значення  $k$ .

Алгоритм 2 або алгоритм LDC — це новий алгоритм кластеризації. Основні ідеї в наступному: (1) обчислення локальної щільності значення всіх даних і впорядкування їх у порядку спадання, (2) вибір значення з високою локальною щільністю як центру кластеризації та (3) кластеризація нецентральної точки даних до найближчого класу. Алгоритм LDC визначається кількома параметрами:

- локальною щільністю

$$\rho_i = \sum_i X(d_{ij} - d_c). \quad (1)$$

$$X(i) = \begin{cases} 1 & i < 0 \\ 0 & i \geq 0 \end{cases}, \quad (2)$$

де  $d_{ij}$  представляє відстань між точками даних  $i$  та  $j$ ;  $d_c$  — це відстань відсікання і є регульованим параметром;  $\rho_i$  представляє кількість точок даних у діапазоні з  $i$  як центр та  $d_c$  як радіус (алгоритм LDC більш чутливий до значення  $\rho_i$ , тому значення відстані зрізу  $d_c$  має великий вплив на ефект кластеризації);

- мінімальна відстань високої щільності

$$\delta_i = \min_{j: \rho_j > \rho_i} d_{ij}, \quad (3)$$

де  $d_{ij}$  представляє відстань між точками даних  $i$  та  $j$ , мінімальна відстань високої щільності  $\delta_i$  представляє собою відстань від точки даних  $i$  до найближчої точки  $j$  серед усіх точок з локальною щільністю, більшою за точку даних  $i$ . Точка з найвищою локальною щільністю, визначається як:

$$\delta_i = \max(d_{ji}). \quad (4)$$

Розглянемо приклад для аналізу LDC алгоритму в деталях:

(1) Діаграма прийняття рішень: LDC — це двовимірний алгоритм. Алгоритм виконує кластеризацію відповідно на рис. 3. На рис. 3(a) показано діаграму розподілу даних, а на рис. 3(b) показано рішення діаграми даних за рис. 3(a). Етикетки точки вибірки розташовані в порядку спадання відповідно до місцевої щільності. Різні кольори представляють різні кластери, а чорні крапки — шум точки.

Судження цього центру кластеризації в основному людськими очима. З діаграми рішень, легко виявити, що точки 1 і точки 10 явно відрізняються від інших точок даних, і вони мають відносно велику локальну щільність  $\rho$  і відносно велику мінімальну відстань високої щільності  $\delta$ . Отже, вони використовуються як центр кластера при кластеризації. Хоча точки 26, 27 і 28 мають більшу мінімальну відстань високої щільності  $\delta$ , локальна щільність  $\rho$  менше. Отже, вони виділяються як точки шуму.

(2) Класифікація точок даних: після визначення центральної точки кластера на діаграмі рішень решта даних кластеризується до центральної точки, найближчої до себе. Алгоритм простий і зрозумілий, має низьку складність і не вимагає ітерацій. Після того, як алгоритм встановлює центральну точку, безпосередньо виконується кластеризація.

(3) Фільтрація шуму: точки шуму впливатимуть на ефект кластеризації. Алгоритм видаляє шумові точки за граничною щільністю. Якщо локальна щільність  $\rho$  даних задовольняє  $\rho_i \leq \mu(\rho) - 2\sigma(\rho)$ , вона визначається як точка шуму та видаляється з набору даних.

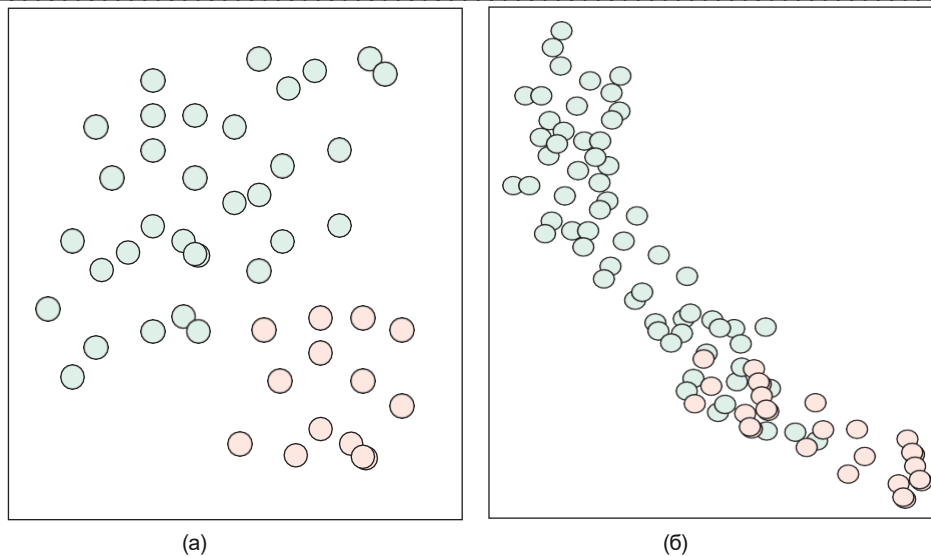


Рис. 3 : Діаграма розподілу даних і діаграма рішень. (а) Діаграма розподілу точок даних. (б) Діаграма прийняття рішень.

#### 4. Оптимізація алгоритму для покращення здатності обробки масивної інформації даних.

Націлюючись на проблему, що результати кластеризації не можуть досягти глобального оптимального значення, спричиненого випадковим вибором початкової центральної точки  $K$ -середніх і впливом точок шуму на кластеризацію, у цій статті пропонується алгоритм LDCK і проводяться експерименти кластеризації, щоб перевірити його ефективність. Алгоритм  $K$ -середніх також має такі недоліки: якість кластеризації чутлива до випадкового вибору початкових точок, вибору  $k$  слід вказати значення, кластеризація також більш чутлива до наявних точок шуму. Далі ця стаття детально аналізує ці недоліки, потім виконує відповідні оптимізації та, нарешті, розпаралелює оптимізовані алгоритми, щоб покращити здатність обробляти великомасштабні дані.

Визначимо ще декілька параметрів для подальших досліджень.

Ступінь міжкластерної дисперсії

$$Disp = \frac{\sum_{i=1}^k Disp_i}{k} \quad (5)$$

$Disp$  представляє ступінь міжкластерної дисперсії,  $Disp_i$  визначає положення центральної точки  $i$ -го кластера, і  $k$  вивчає кількість кластерів.

Ступінь внутрішньокластерної агрегації

$$Aggr = \frac{\sum_{i=1}^k Aggr_i}{k} \quad (6)$$

$Aggr$  представляє ступінь внутрішньокластерної агрегації,  $k$  представляє кількість кластерів, а  $Aggr_i$  являє собою середню відстань від усіх елементів  $i$ -го кластера до центру.

Значення оцінки кластеризації

$$E = \frac{Aggr_k - Aggr_{k-1}}{Disp_k - Disp_{k-1}} \quad (7)$$

де  $Disp_k$  —  $k$ -й ступінь міжкластерної дисперсії,  $Disp_{k-1}$  визначає  $k$ -й ступінь міжкластерної дисперсії,  $Aggr_k$  означає  $k$ -й ступінь внутрішньокластерної агрегації,  $Aggr_{k-1}$  визначає  $k - 1$ -й ступінь внутрішньокластерної агрегації, і  $E$  — значення оцінки кластера.

Для визначення ефективності розглянемо приклад. Взявши дані з бази даних вибираємо два типи даних А та В (властивості функціональної стійкості) як набір експериментальних даних. Їх обсяги даних становлять 148 і 181, відповідно, атрибути даних дорівнюють 4 і 3



відповідно, і всі типи даних всіх 3.

Щоб також показати відмінності оціночних значень  $E$  як  $\log_2^E - \log_2^{E_0}$  у процесі кластеризації А та В, а лінійний графік малюється, як показано на рис. 4.

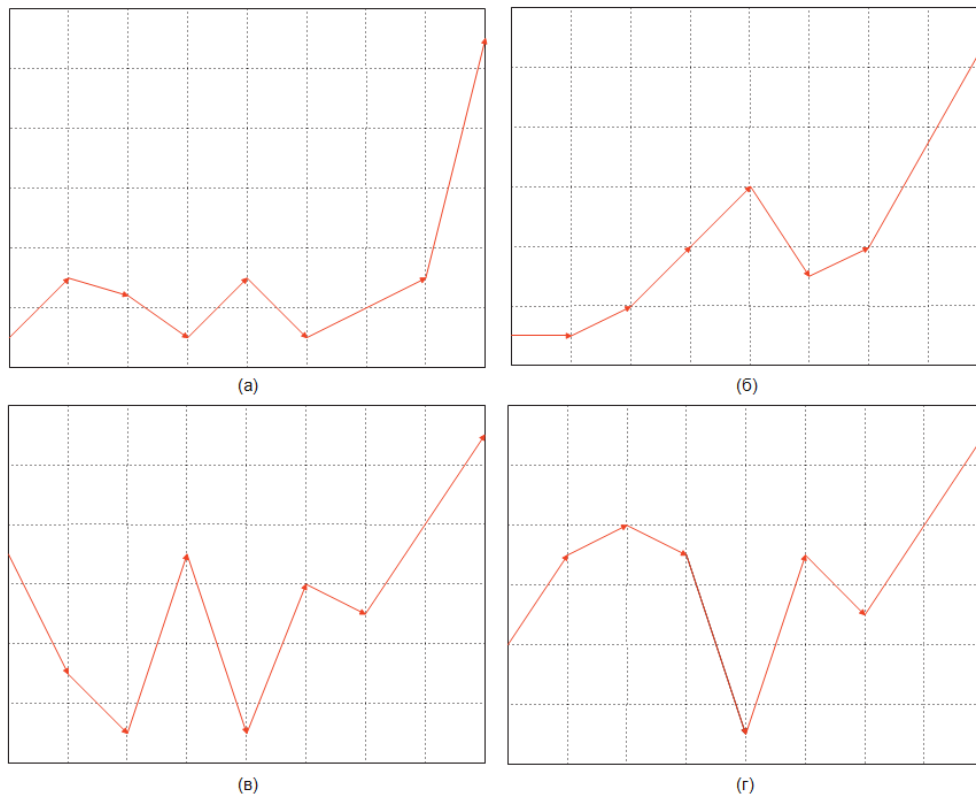


Рис. 4: Лінійна діаграма оціночних значень  $E$  і  $\log_2^E - \log_2^{E_0}$  А та В. (а) Зміна оціночного значення  $E$  для А. (б) Зміна оцінки  $E$  для В. (в) Зміна оцінки А значення  $\log_2^E - \log_2^{E_0}$ . (г) Зміна оцінки В значення  $\log_2^E - \log_2^{E_0}$ .

На рис. 4, коли значення  $k$  більше 2, значення  $E$  має невелику зміну, а значення  $\log_2^E - \log_2^{E_0}$  менше ніж 1, який вказує на те, що алгоритм може продовжувати ітерацію як і поєднання розумне. При цьому значення  $k$  змінюється від 3 до 2 значення  $E$  раптово зростає, а значення  $\log_2^E - \log_2^{E_0}$  більше за 1. Це означає, що злиття є непотрібним, тому кількість кластерів в обох наборах даних становить 3. Крім того, ці експериментальні результати відповідають дійсному числу 3 із двох наборів даних. Це демонструє ефективність запропонованого алгоритму.

## 5. Висновок

Вибір найкращого алгоритму за умов конкретного завдання має обґрунтовано здійснюватися особою, яка приймає рішення, оскільки вкрай важливим є процеси інтерпретації та оцінювання здобутих результатів після їх аналізу. На даній стадії головну роль відіграє експерт у досліджуваній предметній галузі, який крім використання критеріїв може на основі апріорних уявлень і знань ключових цільових показників здійснити додаткову верифікацію результатів для подальшого прийняття рішень. Для досягнення максимального результату потрібен комплексний підхід до аналізу даних, що включає в себе як використання апріорних знань фахівців для попереднього оброблення даних і інтерпретації результатів, так і застосування спеціалізованих алгоритмів.

У роботі проведено аналіз недоліків, виконано оптимізацію та розпаралелювання оптимізованих алгоритмів, щоб покращити здатність обробляти великий масив даних та підвищити ефект від діяльності інтелектуальної інформаційної системи закладу вищої освіти за допомогою використання хмарних обчислень та машинного навчання інтелектуальної інформаційної системи.

**Список використаної літератури:**

1. Sobchuk V., Pichkur V. On conditions for ensuring of functional stability of information systems in a manufacturing enterprise // Abstracts of XIX International Scientific and Practical Conference. Brussels, Belgium. April 08-09, 2021. P. 219-221.
2. Zamrii I. V. Strategy of management of functional stability of the information system of the industrial enterprise // International Journal of Progressive Sciences and Technologies, Vol. 29, No 1 October 2021. ISSN: 2509-0119. pp. 659-667.
3. Собчук В.В., Мусієнко А.П., Ільїн О.Ю. Аналіз використання ієрархічної структури для забезпечення функціональної стійкості автоматизованої системи управління підприємством // Науковий журнал «Телекомунікаційні та інформаційні технології». К.: ДУТ, 2018. № 4 (61). С. 53 – 61.
4. G. J. Joyia, R. M. Liaqat, and A. Farooq, “Internet of Medical Things (IOMT): applications, benefits and future challenges in healthcare domain,” *Journal of Communication*, vol.12, no.4, pp. 240–247, 2017.
5. N. Kshetri, “The evolution of the internet of things industry and market in China: an interplay of institutions, demands and supply,” *Telecommunications Policy*, vol. 41, no. 1, pp. 49–67, 2017.
6. S.Siboni, V.Sachidananda, Y.Meidanetal., “Security testbed for internet-of-things devices,” *IEEE Transactions on Reliability*, vol. 68, no. 1, pp. 23–44, 2019.
7. Y. Yang, M. Zhong, H. Yao, F. Yu, X. Fu, and O. Postolache, “Internet of things for smart ports: technologies and challenges,” *IEEE Instrumentation and Measurement Magazine*, vol. 21, no. 1, pp. 34–43, 2018.
8. A. Heiskanen, “The technology of trust: how the internet of things and blockchain could usher in a new era of construction productivity,” *Construction Research and Innovation*, vol. 8, no. 2, pp. 66–70, 2017.
9. M. Wolf and D. Serpanos, “Safety and security in cyberphysical systems and internet-of-things systems,” *Proceedings of the IEEE*, vol. 106, no. 1, pp. 9–20, 2017.
10. Собчук В.В., Замрій І.В., Барабаш О.В., Мусієнко А.П. Методологія побудови функціонально стійкої інтелектуальної інформаційної системи виробничого підприємства // Вісник Київського національного університету імені Тараса Шевченка. Серія фізико-математичні науки, 2021, №.4, с. 116-127.
11. Zamrii I. V. Strategy of management of functional stability of the information system of the industrial enterprise // International Journal of Progressive Sciences and Technologies, Vol. 29, No 1 October 2021. ISSN: 2509-0119. pp. 659-667.
12. Valentyn Sobchuk, Iryna Zamrii, and Serhii Laptiev. Ensuring Functional Stability of Technological Processes as Cyberphysical Systems Using Neural Networks. Springer Nature Switzerland AG 2023 O. Arsenyeva et al. (Eds.): Smart Technologies in Urban Engineering Proceedings of STUE-2022. Lecture Notes in Networks and Systems, Volume 536. pp. 581–592, 2023.
13. Замрій І.В., Собчук А.В., Лаптев С.О., Лаптева Т.О., Копитко С.Б. Алгоритм контролю та прогнозування функціональної стійкості складних інформаційно-технічних систем // Телекомунікаційні та інформаційні технології, 2022, №1 (74). С. 4-14.

**References**

1. Sobchuk V., Pichkur V. On conditions for ensuring of functional stability of information systems in a manufacturing enterprise // Abstracts of XIX International Scientific and Practical Conference. Brussels, Belgium. April 08-09, 2021. P. 219-221.
2. Zamrii I. V. Strategy of management of functional stability of the information system of the industrial enterprise // International Journal of Progressive Sciences and Technologies, Vol. 29, No 1 October 2021. ISSN: 2509-0119. pp. 659-667.

3. Sobchuk V.V., Musienko A.P., Ilyin O.Y. Analysis of the use of a hierarchical structure to ensure the functional stability of the automated enterprise management system // Scientific journal "Telecommunication and information technologies". K. : DUT, 2018. No 4 (61). P. 53 - 61.
4. G. J. Joyia, R. M. Liaqat, and A. Farooq, "Internet of Medical Things (IOMT): applications, benefits and future challenges in healthcare domain," *Journal of Communication*, vol.12,no.4, pp. 240–247, 2017.
5. N. Kshetri, "The evolution of the internet of things industry and market in China: an interplay of institutions, demands and supply," *Telecommunications Policy*, vol. 41, no. 1, pp. 49–67, 2017.
6. S.Siboni,V.Sachidananda,Y.Meidanetal.,"Security testbed for internet-of-things devices," *IEEE Transactions on Reliability*, vol. 68, no. 1, pp. 23–44, 2019.
7. Y. Yang, M. Zhong, H. Yao, F. Yu, X. Fu, and O. Postolache, "Internet of things for smart ports: technologies and challenges," *IEEE Instrumentation and Measurement Magazine*, vol. 21, no. 1, pp. 34–43, 2018.
8. A. Heiskanen, "The technology of trust: how the internet of things and blockchain could usher in a new era of construction productivity," *Construction Research and Innovation*, vol. 8, no. 2, pp. 66–70, 2017.
9. M. Wolf and D. Serpanos, "Safety and security in cyberphysical systems and internet-of-things systems," *Proceedings of the IEEE*, vol. 106, no. 1, pp. 9–20, 2017.
10. Sobchuk V.V., Zamrii I.V., Barabash O.V., Musienko A.P. Methodology of building a functionally stable intellectual information system of a production enterprise // *Bulletin of Taras Shevchenko Kyiv National University. Series of physical and mathematical sciences*, 2021, No. 4, c. 116-127.
11. Zamrii I. V. Strategy of management of functional stability of the information system of the industrial enterprise // *International Journal of Progressive Sciences and Technologies*, Vol. 29, No 1 October 2021. ISSN: 2509-0119. pp. 659-667.
12. Valentyn Sobchuk, Iryna Zamrii, and Serhii Laptiev. Ensuring Functional Stability of Technological Processes as Cyberphysical Systems Using Neural Networks. Springer Nature Switzerland AG 2023 O. Arsenyeva et al. (Eds.): Smart Technologies in Urban Engineering Proceedings of STUE-2022. Lecture Notes in Networks and Systems, Volume 536. pp. 581–592, 2023.
13. Zamrii I.V., Sobchuk A.V., Laptev S.O., Lapteva T.O., Kopytko S.B. Algorithm of control and prediction of functional stability of complex information and technical systems // *Telecommunications and information technologies*, 2022, No. 1 (74). P. 4-14.