

С. М. ШЕВЧЕНКО

кандидат педагогічних наук, доцент,
доцент кафедри інформаційної та кібернетичної безпеки
імені професора Володимира Бурячка
Київський університет імені Бориса Грінченка
ORCID: 0000-0002-9736-8623

Ю. Д. ЖДАНОВА

кандидат фізико-математичних наук, доцент,
доцент кафедри інформаційної та кібернетичної безпеки
імені професора Володимира Бурячка
Київський університет імені Бориса Грінченка
ORCID: 0000-0002-9277-4972

Т. І. ШЕВЦОВА

магістрант кафедри інженерії програмного забезпечення
Державний університет інформаційно-комунікаційних технологій
ORCID: 0009-0007-2802-8127

ЗАСТОСУВАННЯ КЛАСТЕРНОГО АНАЛІЗУ ДЛЯ ПРОСУВАННЯ БІЗНЕСУ У СОЦІАЛЬНИХ МЕРЕЖАХ

Використання соціальних мереж для досягнення бізнес-цілей, зокрема просування своїх послуг (товарів), є затребуваним у сучасному світі. У 2022 році понад 4,59 мільярдів людей (Statista) користувалися соціальними мережами. Враховуючи таку колосальну аудиторію активних користувачів, маркетологи вбачають важливим та значущим використання соціальних мереж для бізнес-діяльності своєї компанії. Це спонукало до дослідження методів та інструментарію просування продукції в Internet.

Огляд наукової та методичної літератури дозволив виділити існуючі методи просування бізнесу: Owned Media; Paid Media; Earned Media; Social Media, основним недоліком яких є негативні коментарі користувачів, що шкодить репутації компанії. У даній статті розглядається можливість застосування методів кластерного аналізу для просування бізнесу у соціальних мережах. Визначено поняття кластерний аналіз, описано найпопулярніші методи кластерного аналізу, представлено типовий механізм його проведення. За основу у дослідженні запропоновані блок-схеми проведення кластеризації методом ближнього сусіда та методом k-means. Виявлені у кожному з них переваги та недоліки.

В якості прикладу проведено розподіл споживачів послуг ТОВ «ЕПАМ СИСТЕМЗ», що надається за одну добу, в залежності від попиту на послуги компанії з урахуванням вікового критерію. Результати показали, які послуги і для якого віку є більш привабливими, що допоможе ефективніше проводити заходи стимулювання обсягів продажів IT-послуг у соціальних мережах, що призведе до зростання обсягів прибутку організації у майбутньому. Результати дослідження можна впровадити в навчальний процес студентів економічного профілю та студентів галузі ІТ Інформаційні технології.

Ключові слова: соціальні мережі, система просування, потреби споживачів послуг, кластерний аналіз, алгоритм найближчого сусіда, алгоритм k-means.

S. M. SHEVCHENKO

PhD, Associate Professor,
Associate Professor at the Department of Information and Cyber Security
named after Professor Volodymyr Buriachok
Borys Grinchenko Kyiv University
ORCID: 0000-0002-9736-8623

YU. D. ZHDANOVA

PhD, Associate Professor,
Associate Professor at the Department of Information and Cyber Security
named after Professor Volodymyr Buriachok
Borys Grinchenko Kyiv University
ORCID: 0000-0002-9277-4972

T. I. SHEVTSOVA

Master's Student at the Department of Software Engineering
National University of Information and Communication Technologies
ORCID: 0009-0007-2802-8127

THE RESEARCH OF CLUSTER ANALYSIS WAYS OF APPLICATION FOR BUSINESS PROMOTION IN SOCIAL NETWORKS

The use of social networks to achieve business goals, in particular to promote one's services (goods), is in demand in today's world. In 2022, more than 4.59 billion people (Statista) used social networks. Given such a colossal audience of active users, marketers consider it important and meaningful to use social networks for their company's business activities. Studying methods and tools to promote products on the Internet is a result of this.

The review of scientific and methodical literature has highlighted existing methods of business promotion: Owned Media; Paid Media; Earned Media; Social Media, the main drawback of which are negative comments from users, which harms the company's reputation. This article considers the possibility of applying methods of cluster analysis to promote business in social networks. The concept of cluster analysis has been defined, the most popular methods of cluster analysis have been described, a typical mechanism for its implementation has been presented. The study is based on block-schemes of clustering by near neighbor method and k-means method. Advantages and disadvantages have been identified in each of them.

As an example, the distribution of consumers of services of «EPAM SYSTEMZ» LLC, which is provided for a day, depending on the demand for services of the company, taking into account the age criterion. The results showed which services and for what age are more attractive, which will help to more effectively carry out measures to stimulate sales of IT services in social networks, which will lead to an increase in the volume of profit of the organization in the future. The results of the study can be introduced into the educational process of students of economic profile and students of branch of knowledge 12 Information Technologies.

Key words: *Keywords: social networks, promotion system, needs of service consumers, cluster analysis, nearest neighbor algorithm, k-means algorithm.*

Постановка проблеми

Сучасний час характеризується різноманітними інноваціями та сучасними технологіями, зокрема соціальними мережами, які можна вважати невід'ємною частиною повсякденного життя. Майже в усіх частинах світу незліченна кількість людей щодня користуються соціальними мережами. Сьогодні практично кожна людина має можливість спілкуватися з іншими людьми за рахунок використання соціальних мереж.

Соціальні мережі застосовуються в якості засобу спілкування та об'єднання людей у всьому світі, що дозволяє людям дізнаватися про нові звичаї та культури, а також сприяє масовому поширенню інформації серед людей (обмін відео та фотографіями, пропозиціями роботи, відгуками тощо). Останнім часом вони знайшли широке застосування в господарській діяльності суб'єктів господарювання. Вплив соціальних мереж проникає в усі аспекти людського життя, як особистого, так і професійного. Суб'єкти господарювання намагаються використовувати потужність соціальних мереж переважно для просування своєї продукції, створення бренду або залучення нових клієнтів. Соціальні мережі допомагають у створенні лояльності існуючих клієнтів до такої міри, що вони це зроблять добровільно та безкоштовно поширюватимуть дані про діяльність певних компаній.

Переваги побудови соціальних мереж як інструменту побудови зв'язків з громадськістю для компанії полягають у тому, що соціальні мережі можуть забезпечити поширення рекламного контенту серед потенційних клієнтів. Крім того, соціальні мережі можуть допомогти налаштувати зворотній зв'язок систем, створювати чати в соціальних мережах і можуть дозволити обмінюватися враженнями від спожитих товарів або послуг. Компанії можуть створювати бренд компанії або бренд продукту за рахунок створення профілів у мережах Facebook і LinkedIn або облікового запису у Twitter. Станом на січень 2023 року Facebook був найбільш часто використовуваною платформою соціальних мереж серед маркетологів у всьому світі. Згідно з глобальним опитуванням, 89 відсотків маркетологів соціальних мереж використовували мережу для просування свого бізнесу, тоді як ще 80 відсотків робили це через Instagram [1]. Оскільки розвиток інформаційних технологій сприяє розвитку віртуальної комунікації, то компанії, що виробляють товари або надають послуги, мають розмовляти зі своїми потенційними споживачами їх мовою.

Аналіз останніх досліджень і публікацій

Обрана проблематика розглядалася у працях таких науковців, як П. Тівасінг, Р. П. Кайперс, Г. Ертуг, Дж. Кантвелл, А. Захір, М. Кілдафф та у працях інших дослідників. Слід підкреслити, що бурхливий розвиток інформаційних технологій змушує компанії адаптуватися до поточних тенденцій у своїй бізнес-діяльності, особливо тих, що стосуються ІТ. Тим більше це стосується суб'єктів господарювання, які здійснюють свою господарську діяльність в онлайн-середовищі. Одним із найбільш використовуваних і яскраво виражених сучасних трендів є соціальні мережі. Саме тому дослідники вивчали, як корисно можна застосовувати соціальні мережі та як це впливає на результативність заходів системи просування товарів чи послуг.

Однією з таких мереж є, наприклад, соціальна платформа Facebook. М. С. Луо розглядала Facebook не лише як сторінку для розміщення повідомлень, а й як платформу, місце, де можна поширити різноманітні програми стимулювання продажів товарів та послуг. Соціальна мережа LinkedIn спочатку була призначена для пошуку людей, яких потрібно залучити до роботи в компанії через знайомих, друзів, а пізніше стала застосовуватися як рекламна платформа [2]. Г. Ертут підкреслював, що сучасні соціальні мережі дозволяють створити професійну ідентичність в Інтернеті та знайти і отримати пропозиції про роботу, бізнес-співпрацю, пропозиції брати участь у виграшах, лотереях, акціях. LinkedIn надавав численні переваги, які можна використовувати для професійного розвитку та початку бажаної кар'єри. Це дозволяло створити професійну анкету, завантажити CV та інформацію про освіту та досвід роботи і залучити потенційних роботодавців та отримати бажану роботу. З іншого боку, це дозволяло формувати бази даних потенційних покупців певного товару або певної послуги [3].

Соціальна мережа Twitter була створена в 2006 (з липня 2023 йде оновлення до X) році і на сьогодні працює на основі так званих мікроблогів. Мікроблоги – короткі повідомлення, довжина яких не перевищує ста сорока символів. Ця мережа, в першу чергу, була призначена для подорожуючих людей, щоб вони могли поділитися своїм досвідом за рахунок використання мобільних телефонів чи інших пристроїв. Р. Равіна-Ріполь визначала, що популярною соціальною мережею була Instagram, безкоштовна програма для обміну фотографіями та відео, доступною для пристроїв iPhone та Android. Люди можуть використовувати Instagram, щоб поширювати фотографії та відео зі своїми підписниками або обраними групами друзів, рекламувати товари, послуги. Використовуються, також, так звані хештеги, які дозволяють легше шукати контент з тією ж тематикою, зберегти сторінки з рекламною інформацією, які привернули увагу [4]. Р. Тівасінг підкреслював, що застосування соціальних мереж як методу реалізації прямих продажів, дозволяє забезпечити проведення ефективного спілкування з людьми, що є сучасною базою електронного маркетингу, застосування яких призводить до зростання обсягів прибутку компанії [5]. У дослідженні [6] доведено, що соціальні мережі впливають на ефективність онлайн-бізнесу у різних секторах принаймні на 30%, хоча Twitter та YouTube румунськими компаніями використовуються нечасто.

Таким чином, переваги соціальних мереж у просуванні бізнесу є очевидними. Це спонукало менеджерів компаній ефективно їх використати для формування та розвитку свого бренду, посилення відвідування їх офіційного сайту, що дозволить збільшити кількість активних покупців, як наслідок, збільшиться прибуток. У дослідженні [7] здійснено огляд основних інструментів і технологій Internet-маркетингу, розкрито сутність кожної категорії, наведено рекомендації щодо застосування інструментів компаніями різного масштабу, а також зазначено найпопулярніші Internet-інструменти для просування продукції: Owned Media; Paid Media; Earned Media; Social Media. Один із способів просування бізнесу у соціальних мережах носить назву Social Media Marketing (SMM). Задачі, які можна вирішити за допомогою SMM: залучення відвідувачів на сайт компанії; підвищення впізнаваності вашого бренду; підтримка репутації; підвищення лояльності до компанії; збільшення обсягу продаж; організація майданчика для комунікації з клієнтами, зворотного зв'язку; запуск і супровід нового проєкту [8–11]. Проте одним із недоліків SMM-просування є негативні коментарі, що може привести до негативних наслідків для компанії.

Останнім часом у наукових джерелах обґрунтовується необхідність застосування методів кластерного аналізу для визначення оцінки ефективності управління бізнесом [12, 13]. Так, кластерний аналіз застосовують для сегментації клієнтів на основі схожих характеристик [12]; для ефективного використання ресурсного потенціалу [13]; для розробки продуктів та послугів на основі потреб клієнтів [12]; для визначення ризиків у бізнес-діяльності та інше. Вище викладене підтвердило важливість та актуальність даного дослідження.

Формулювання мети дослідження

Метою дослідження є обґрунтування можливостей застосування методів кластерного аналізу для просування бізнесу у соціальних мережах.

Виклад основного матеріалу дослідження

Уперше термін «кластерний аналіз» було введено у 1939 році (Труон Роберт). Проте розквіту цей метод не мав, бо вимагав громіздкі обрахунки в процесі виконання. Лише на початку нашого століття інтенсивний розвиток потужних комп'ютерів відкрив можливість для обробки великих даних з метою їх групування по схожим характеристикам. Кластерний аналіз – метод аналізу даних, який досліджує природні групи в наборі даних, відомих як кластери. Цей метод дає можливість відносити об'єкти до однієї групи не за одним показником, а за декількома водночас. Також допомагає віднайти структуру даних, що неможливо зробити з боку експерта чи зовнішньому аналізу [14, 15]. Кластерному аналізу не потрібно групувати точки даних у будь-які попередньо визначені групи, що означає, що це – метод навчання без контролю. Виокремлюють жорстку та м'яку кластеризацію. Жорстка кластеризація передбачає, що кожна точка вхідних даних частково або повністю належить до кластеру або ні. М'яка кластеризація визначає, що замість того, щоб поміщати кожен вхідну точку даних в окремий кластер, призначається ймовірність перебування цієї точки даних у цих кластерах.

Відомо понад ста алгоритмів кластеризації. Розглянемо деякі з них докладніше. Ієрархічні методи діляться на два види: агломеративні та дивизимні. Перші базуються на тому, що усі об'єкти розташовують в окремі класи,

а потім, враховуючи метрику подібності, ці об'єкти склеюють і тим самим зменшують кількість кластерів поки не отримуються один. Другі методи – навпаки: усі об'єкти – один кластер. Із збільшенням відстані їх розподіляють по окремим.

Центроїдні моделі – ітераційні алгоритми кластеризації, у яких поняття подібності виводиться на основі близькості точки даних до центроїда або центру кластерів. Найчастіше використовуються такі з них, як метод k-means та Fuzzy c-means. У цих моделях слід задати кількість кластерів заздалегідь. Ці моделі виконуються ітеративно, щоб знайти локальний оптимум даних.

Моделі щільності шукають у просторі даних області з різною щільністю точок даних у просторі даних. Вони виділяють різні щільні регіони та призначають точки даних у цих регіонах одному кластеру. Популярними прикладами моделей щільності є DBSCAN і OPTICS. Ці моделі є особливо корисними для ідентифікації кластерів довільної форми та виявлення викидів, оскільки вони можуть виявляти та відокремлювати точки, розташовані в розріджених областях простору даних, а також точки, які належать до щільних областей.

Класифікацію алгоритмів кластеризації відображено на рис. 1.

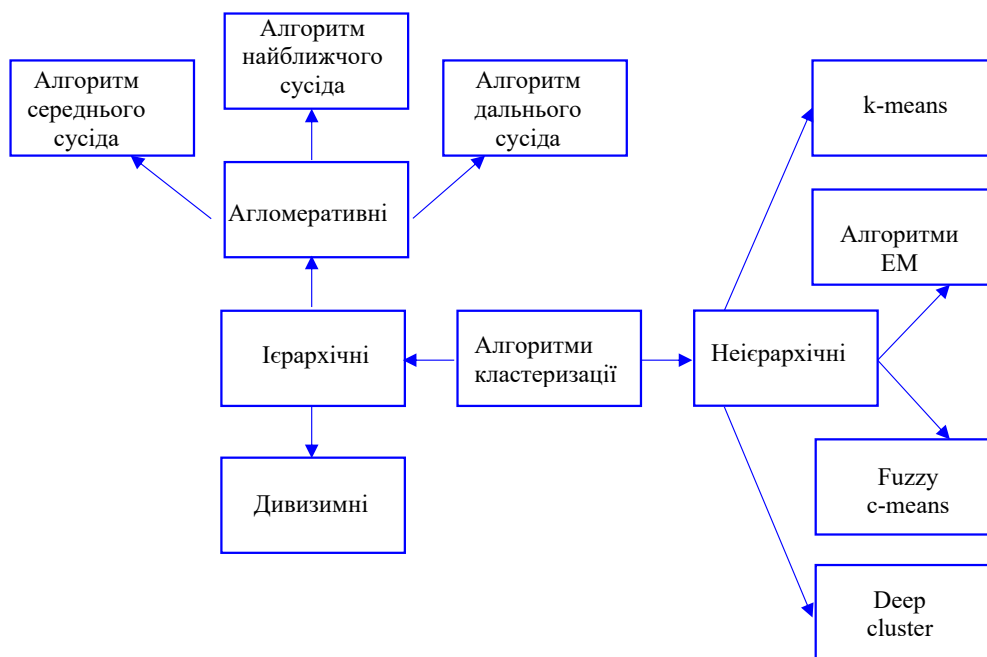


Рис. 1. Найбільш популярні алгоритми кластеризації

Кластеризація дозволяє поєднати в собі такі методи, як методи розбиття, ієрархічної кластеризації, нечіткої кластеризації, реалізація кластеризації на основі щільності та на основі моделі. У кластеризації на основі щільності дані групуються за областями високої концентрації точок даних, оточених областями низької концентрації точок даних. В основному алгоритм знаходить місця, у яких багато точок даних, і викликає ці кластери. Кластери можуть мати будь-яку форму, вони не є обмеженими очікуваними умовами. Алгоритми кластеризації цього типу не намагаються призначити кластерам викиди, тому вони ігноруються.

Методи кластеризації на основі розподілу всі точки даних вважаються частинами кластера на основі ймовірності того, що вони належать до певного кластера. Існує центр-точка, і зі збільшенням відстані точки даних від центру ймовірність того, що вона буде частиною цього кластера, зменшується. Кластеризація на основі центроїда є чутливою до початкових параметрів, але є швидким і ефективним методом аналізу даних. Ці типи алгоритмів розділяють точки даних на основі кількох центроїдів у даних. Кожна точка даних призначається кластеру на основі квадрата її відстані від центроїда. Це – найбільш часто використовуваний тип кластеризації. Кластеризація на основі ієрархії зазвичай використовується для ієрархічних даних, які отримуються з бази даних компанії або таксономії. Це дозволяє створити дерево кластерів, щоб все було організовано за принципом побудови ієрархії зверху-вниз.

Аналіз літератури з досліджуваної теми [12, 13, 16–19] дозволив виділити типовий алгоритм виконання кластерного аналізу у різних галузях для різних об'єктів, зокрема у бізнесі.

Використаємо результати дослідження [16]. Нехай множина $A = \{a_1, a_2, \dots, a_n\}$ – множина об'єктів, B – множина номерів кластерів. Вибирається метрика (найчастіше формула відстані) $d_{ij}(x_{ik}, x_{jk})$. Необхідно розбити

множину A на підмножини (кластери), які не перетинаються, і кожен кластер містить об'єкти близькі за метрикою $d_{ij}(x_{ik}, x_{jk})$, а об'єкти різних класів істотно відрізнялися. Кожному об'єкту a_i приписується номер кластера B_i .

Множина A може складатися із об'єктів, які мають різні одиниці вимірювання або різний діапазон представлених значень, тому потрібно здійснити нормування вхідних даних. При кластерному аналізі є два основні способи нормалізації даних: MinMax-нормалізація та Z-нормалізація.

MinMax-нормалізація здійснюється наступним чином:

$$x' = \frac{x - \min[X]}{\max[X] - \min[X]} \quad (1)$$

У разі всі значення будуть у діапазоні від 0 до 1; дискретні бінарні значення визначаються як 0 та 1.

Z-нормалізація:

$$x' = \frac{x - M[X]}{\sigma[X]}, \quad (2)$$

де $M[X]$ – математичне сподівання, $\sigma[X]$ – середнє квадратичне відхилення.

Введемо метрику для оцінки досліджуваних об'єктів і відстаней між ними. Як правило, відстань між двома об'єктами представлена невід'ємною функцією близькості, яка вводиться для будь-яких об'єктів кластерного аналізу. Якщо розглядати реальні умови контролю, то ефективніше порівнювати об'єкти за інтегральними характеристиками. На жаль, цей спосіб далеко не завжди застосовний через неможливість узгодження всіх одиниць виміру з урахуванням різниці метричних полів.

У кластерному аналізі можуть використовуватися міри подібності: коефіцієнти кореляції, міри відстані, коефіцієнти асоціативності, ймовірнісні коефіцієнти подібності. За міру подібності у нашому дослідженні будемо використовувати евклідову відстань:

$$d_{ab} = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2} \quad (3)$$

Надалі вибирається і впроваджується у процес відповідний алгоритм кластерного аналізу та у результаті здійснити перевірку достовірності результатів.

Метою нашої експериментальної роботи є визначення окремих груп на основі вибору продуктів, запропонованих компанією ТОВ «ЕПАМ СИСТЕМЗ». Використаємо метод найближчого сусіда та метод k-means.

Перевага алгоритму найближчого сусіда, що відноситься до ієрархічного алгоритму кластеризації полягає у тому, що є наочний результат, простота реалізації, використовується до широкого кола сфер, являється пошуком найкращого рішення із можливих.

Недоліки найближчого сусіда: зберігає всю вибірку об'єктів, що провокує до витрат пам'яті, якщо серед об'єктів існує викид (тобто, об'єкт розташований всередині чужого класу), то всі об'єкти, які будуть знаходитись найближче до всіх інших, будуть класифікуватись неправильно.

Алгоритм виконання даного методу представлено на рисунку 2.

Переваги алгоритму k-means, що відноситься до неієрархічних алгоритмів кластеризації, – це зрозумілість та швидкість виконання, можливість перевірки статистичної значимості відмінностей між виділеними кластерами. Недоліки: потрібно наперед задати кількість кластерів для розбиття, результат роботи залежить від того, яким чином були визначені початкові центри кластерів.

Процес виконання алгоритму k-means представлено на рисунку 3 [17].

Задамо k кластерів.

1. Обчисліть відстань від кожної точки даних d_i ($1 \leq i \leq n$) до всіх центроїдів c_j ($1 \leq j \leq k$) як $d(d_i, c_j)$.
2. Для кожної точки даних d_i знайдіть найближчий центроїд c_j і призначте d_i кластеру j .
3. Встановити ClusterId[i]=j; // j:Id найближчого кластера.
4. Встановити Nearest_Dist[i]= $d(d_i, c_j)$.
5. Для кожного кластера j ($1 \leq j \leq k$) перерахуйте центроїди.
6. Повторіть.
7. Для кожної точки даних d_i :
 - 7.1. Обчисліть його відстань від центроїда поточного найближчого кластера.
 - 7.2. Якщо ця відстань менша або дорівнює поточній найближчій відстані, точка даних залишається в кластері.

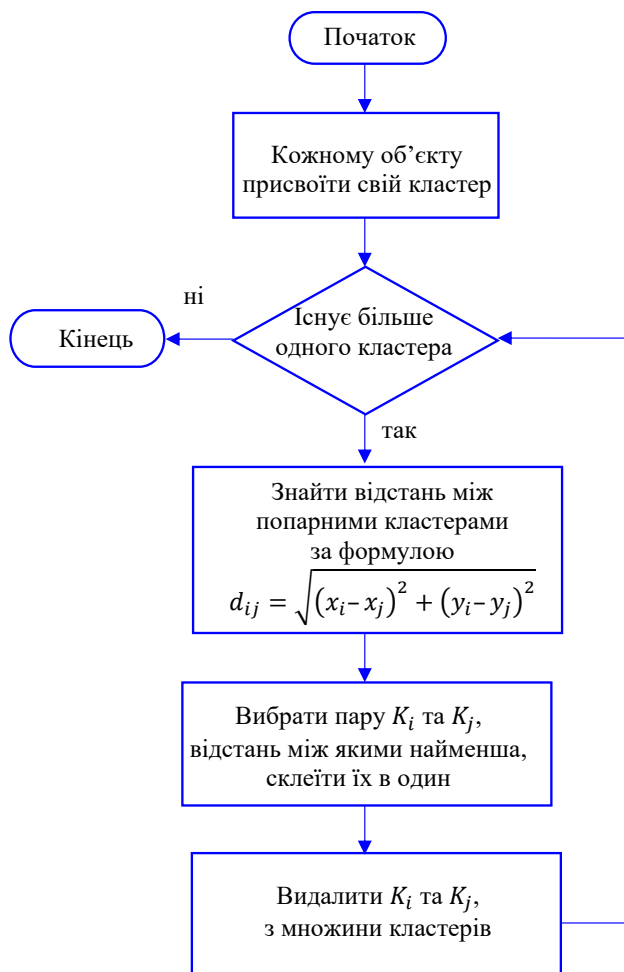


Рис. 2. Блок-схема алгоритма кластеризації методом найближчого сусіда

Інакше

7.2.1. Для кожного центроїда c_j ($1 \leq j \leq k$) обчисліть відстань $d(d_i, c_j)$.

7.2.2. Призначте точку даних d_i кластеру з найближчий центроїд c_j .

7.2.3. Встановити $\text{ClusterId}[i]=j$.

7.2.4. Встановити $\text{Nearest_Dist}[i]=d(d_i, c_j)$.

8. Для кожного кластера j ($1 \leq j \leq k$) перерахуйте центроїди, цей процес повторюється до тих пір, поки центроїди не відокремляться суттєво.

ТОВ «ЕПАМ СИСТЕМЗ» входить до найбільших ІТ-компаній світу, будучи найбільшим виробником замовного програмного забезпечення та бізнес-додатків. До основних видів діяльності ТОВ «ЕПАМ СИСТЕМЗ» належать:

- 1 – курси з програмування,
- 2 – навчання курсантів (магістерський освітній рівень),
- 3 – створення програмного продукту,
- 4 – консультування з питань інформатизації, надання програмного забезпечення,
- 5 – діяльність у сфері інформаційних технологій і комп'ютерних систем,
- 6 – оброблення та розміщення даних.

Проведемо кластерний аналіз шести послуг, що надає ТОВ «ЕПАМ СИСТЕМЗ» своїм клієнтам за одну добу, враховуючи поділ споживачів на дві вікові групи: від 18 до 40 років та від 41 до 60 років. Застосуємо параметр x для позначення попиту на послуги клієнтів ТОВ «ЕПАМ СИСТЕМЗ» вікової групи від 18 до 40 років, а параметр y – для позначення попиту на послуги клієнтів ТОВ «ЕПАМ СИСТЕМЗ» вікової групи від 41 до 60 років. Вихідні дані для кластерного аналізу подано у таблиці 1.

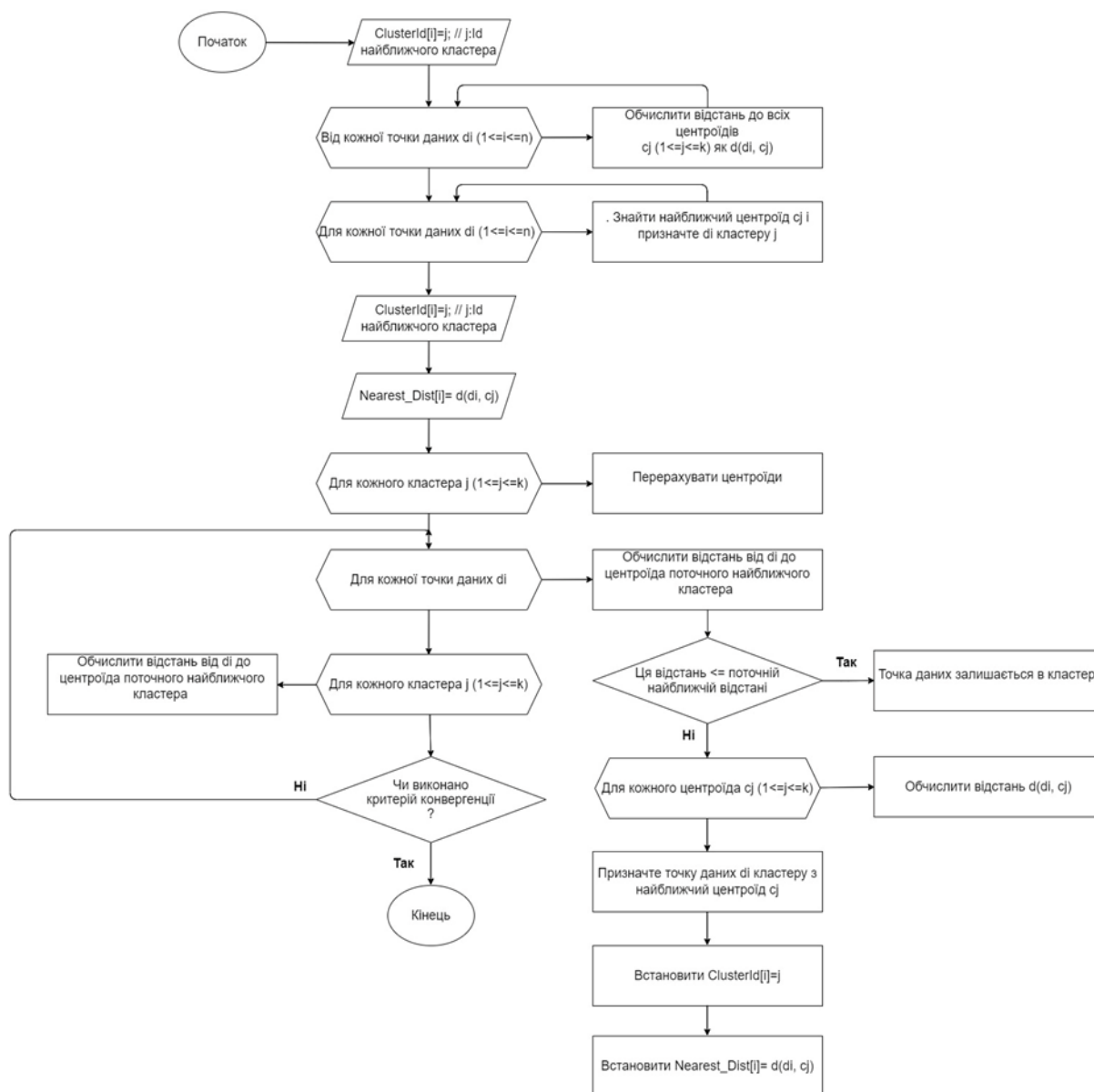


Рис. 3. Блок-схема алгоритма кластеризації методом k-means

Таблиця 1

Вихідні дані для кластерного аналізу

	<i>x</i>	<i>y</i>
1	2	9
2	5	11
3	6	7
4	12	6
5	13	6
6	14	5

У відповідності до формули (3), застосовуючи вихідні дані для кластерного аналізу, вдається розрахувати елементи нової матриці. Результати розрахунків показано у таблицях 2–6.

Таблиця 2

Перший етап кластеризації

	1	2	3	4	5	6
1	0,00	3,61	4,47	10,44	11,40	12,65
2	3,61	0,00	4,12	8,60	9,43	10,82
3	4,47	4,12	0,00	6,08	7,07	8,25
4	10,44	8,60	6,08	0,00	1,00	2,24
5	11,40	9,43	7,07	1,00	0,00	1,41
6	12,65	10,82	8,25	2,24	1,41	0,00

Таблиця 3

Другий етап кластеризації

	1	2	3	4,5	6
1	0,00	3,61	4,47	10,44	12,65
2	3,61	0,00	4,12	8,60	10,82
3	4,47	4,12	0,00	6,08	8,25
4,5	10,44	8,60	6,08	0,00	2,24
6	12,65	10,82	8,25	2,24	0,00

Таблиця 4

Третій етап кластеризації

	1	2	3	4,5,6
1	0,00	3,61	4,47	10,44
2	3,61	0,00	4,12	8,60
3	4,47	4,12	0,00	6,08
4,5,6	10,44	8,60	6,08	0,00

Таблиця 5

Четвертий етап кластеризації

	1,2	3	4,5,6
1,2	0,00	4,12	8,60
3	4,12	0,00	6,08
4,5,6	8,60	6,08	0,00

Таблиця 6

Результат кластеризації

	1,2,3	4,5,6
1,2,3	0,00	6,08
4,5,6	6,08	0,00

З таблиць 2–6 можна зробити висновок, що для застосування для вибірки алгоритму найближчого сусіда знадобилися б чотири етапи. На першому етапі (таблиця 2) серед елементів матриці обираються дві групи елементів, які за значенням є найближчими між собою, що означає основу групування наборів подібних даних. Саме тому обрано об'єкти груп 4 та 5. Елементи групування виділено синім кольором, а елементи, що виділені білим кольором, переносяться у наступну матрицю без змін. Що ж стосується об'єктів, виділених синім кольором, то з двох груп даних у наступну матрицю переносяться найменше з розглянутих значень. На другому етапі (таблиця 3) за вищезазначеним алгоритмом здійснюється поєднання об'єктів груп 4, 5 та 6. На третьому етапі (таблиця 4) реалізується поєднання об'єктів груп 1 та 2 через те, що в обраних групах знаходяться елементи з найменшими значеннями. На четвертому етапі (таблиця 5) здійснюється поєднання об'єктів груп 1, 2 та 3. За результатами кластеризації (таблиця 6) виокремлено об'єкти двох кластерів груп 1, 2, 3 та 4, 5, 6 відстань між якими складає 6,08 умовних одиниць. Таким чином, споживачі віком від 18 до 40 років віддають перевагу отриманню послуг під номерами 1, 2, 3, а споживачі віком від 41 до 60 років – послуг під номерами 4, 5, 6. Візуально це добре видно на дендрограмі (рис. 4).

Саме тому ТОВ «ЕПАМ СИСТЕМЗ», розуміючи що краще запропонувати своїм клієнтам певної вікової групи, допоможе ефективніше проводити заходи стимулювання актуалізації продажів ІТ-послуг у соціальних мережах, що призведе до зростання обсягів прибутку підприємства у майбутньому.

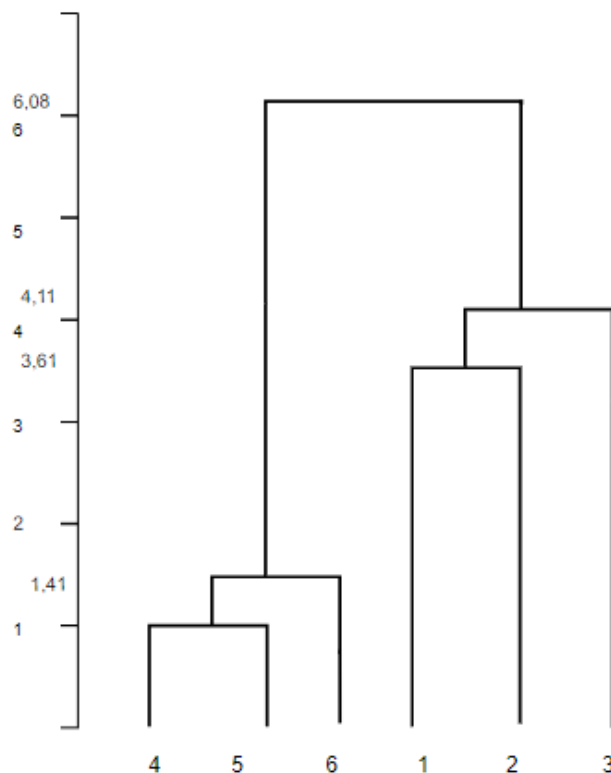


Рис. 4. Дендрограма кластеризації даних компанії ТОВ «ЕПАМ СИСТЕМЗ»

Результат кластеризації даних компанії ТОВ «ЕПАМ СИСТЕМЗ» за методом *k*-means показано на рис. 5. Маємо два згустки (кластери) із своїм центроїдом відповідно. Невеликий обсяг статистичних даних пов’язаний з процесом розуміння застосування цих методів для створення програмного забезпечення.

Наразі створюється програмне забезпечення для розв’язання відповідних завдань.

Висновки

В умовах розвитку інноваційних технологій комунікація в соціальних мережах ефективно застосовується для здійснення просування послуг підприємства в Інтернеті. Саме математичні методи, зокрема кластерний аналіз, можуть забезпечити об’єктивні результати у процесі бізнес-діяльності компанії. Розробка програмного забезпечення надасть можливість здійснювати обчислення дуже швидко, що дозволить керівництву компанії та менеджерам приймати більш обґрунтовані рішення.

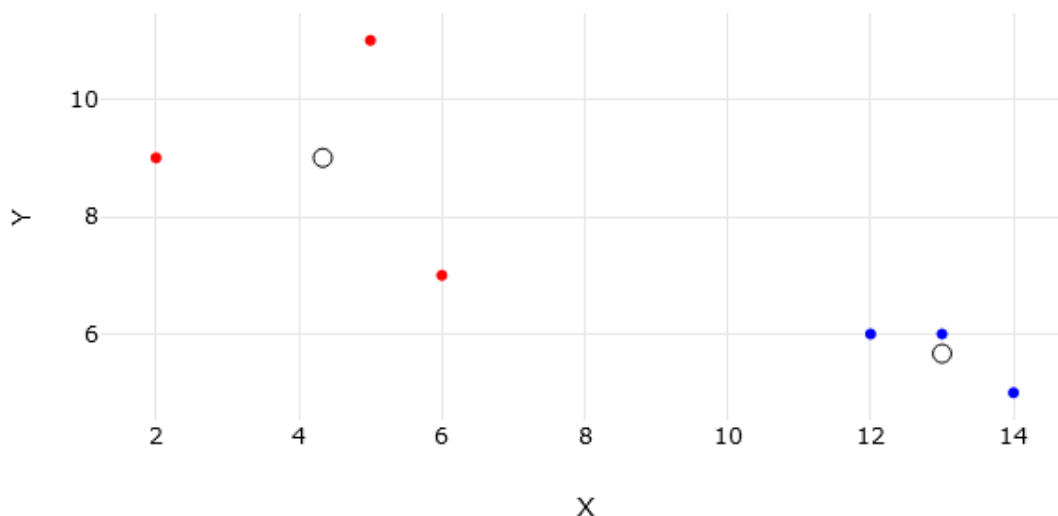


Рис. 5. Результат кластеризації даних компанії ТОВ «ЕПАМ СИСТЕМЗ» за методом *k*-means

Список використаної літератури

1. Leading social media platforms used by marketers worldwide as of January 2023. Statista. (2023) URL: <https://www.statista.com/statistics/259379/social-media-platforms-used-by-marketers-worldwide/>
2. Luo C. Analyzing the impact of social networks and social behavior on electronic business during COVID-19 pandemic. *Information Processing and Management*. №. 58. 2021. P. 37.
3. Cuypers R. P., Ertug G., Cantwell J., Zaheer A., Kilduff M. Making connections: social networks in international business. *Journal of International Business Studies*. №. 51 (5). 2020. P. 714–736.
4. Ahumada-Tello E., Ravina-Ripoll R., Galvez-Albarracin E. G. Social networks and academic performance self-perception in business sciences students. *Social science*. №. 36 (66). 2020. P. 105–117.
5. Tiwasing P. Social media business networks and SME performance: a rural-urban comparative analysis. *Growth and Change*. №. 52 (3). 2021. P. 1892–1913.
6. Alexandra Ioanid, Cezar Scarlat. Factors Influencing Social Networks Use for Business: Twitter and YouTube Analysis. *Procedia Engineering*, Volume 181, 2017, Pages 977–983, ISSN 1877-7058. URL: <https://doi.org/10.1016/j.proeng.2017.02.496>.
7. Ілляшенко С. М., Іванова Т. Є. Інструменти та методи просування продукції в Internet: аналітичний огляд. *Маркетинг і менеджмент інновацій*. № 3. 2015. С. 20–32. URL: http://nbuv.gov.ua/UJRN/Mimi_2015_3_4.
8. Шевчук І. Б. Бізнес у соціальних мережах: Навч. посіб. Львів: Видавництво ННБК «АТБ», 2021. 219 с. URL: https://financial.lnu.edu.ua/wp-content/uploads/2021/11/BSM_navchalnyu-posibnyk_2.pdf
9. Варна М. Що таке SMM і як працює маркетинг у соціальних мережах. Гайд для початківців. (2023) URL: <https://netpeak.net/uk/blog/shcho-take-smm-i-yak-pratsyue-marketing-u-sotsial-nikh-merezhakh-gayd-dlya-pochatktivtiv/>
10. Маркетинг в соціальних мережах. SMM агенство (2022) URL: <https://cases.media/article/marketing-v-socialnikh-merezhakh-smm-agenstvo>
11. Lisa Harris and Alan Rae. Social networks: the future of marketing for small business. *Journal Of Business Strategy*. Vol. 30 No. 5. 2009.
12. Koirala, Jyoti, Understanding the Use of Cluster Analysis in Business (March 27, 2023). URL: <https://ssrn.com/abstract=4400674> or <http://dx.doi.org/10.2139/ssrn.4400674>
13. Безпарточний М. Г. Використання кластерного аналізу при оцінці ефективності діяльності торговельних підприємств. *Торгівля, комерція, підприємництво: збірник наукових праць*. Львів: Львівська комерційна академія. Вип. 17. 2014. С. 24–27.
14. Jain A. K., Dubes R. C. (1988). *Algorithms for Clustering Data*. Upper Saddle River, NJ: Prentice-Hall, Inc. URL: <http://dl.acm.org/citation.cfm?id=SERIES10022.42779>
15. Jain A.K, Murty M.N., Flynn P.J. Data Clustering: A Review. *ACM Computing Surveys*. Vol. 31, No. 3. 1999. P. 264–323. URL: http://www.eecs.northwestern.edu/~yingliu/datamining_papers/survey.pdf
16. Шевченко С.М., Жданова Ю.Д., Негоденко О.В., Куцук В.А. Модель експертної системи для медичного скринінгу на основі методів кластерного аналізу. *Moderní aspekty vědy: XXVII. Díl mezinárodní kolektivní monografie*. Mezinárodní Ekonomický Institut s.r.o. Česká republika: Mezinárodní Ekonomický Institut s.r.o., 2023. С. 478–494. URL: <http://perspectives.pp.ua/public/site/mono/mono-27.pdf>
17. K. A. Abdul Nazeer, M. P. Sebastian. Improving the Accuracy and Efficiency of the k-means Clustering Algorithm. *Proceedings of the World Congress on Engineering*. 2009. Vol I WCE 2009, July 1–3, 2009, London, U.K.
18. Лотиш О.Я. Кластерний аналіз в сегментації галузі. *Вісник Одеського національного університету. Економіка*. Випуск 5 (78). Том 24. 2019. С. 37–42. URL: <https://doi.org/10.32782/2304-0920/5-78-6>
19. Роскладка А.А., Роскладка Н.О., Дзигман О.О. Кластерний аналіз клієнтської бази даних підприємств сфери послуг. *Агроевім*. № 16. 2019. С. 8–17. URL: [https://economics.kntu.kr.ua/pdf/2\(35\)/17.pdf](https://economics.kntu.kr.ua/pdf/2(35)/17.pdf)

References

1. Leading social media platforms used by marketers worldwide as of January 2023. Statista. (2023) URL: <https://www.statista.com/statistics/259379/social-media-platforms-used-by-marketers-worldwide/>
2. Luo C. (2021) Analyzing the impact of social networks and social behavior on electronic business during COVID-19 pandemic. *Information Processing and Management*. №. 58. Pp. 37.
3. Cuypers R. P., Ertug G., Cantwell J., Zaheer A., Kilduff M. (2020) Making connections: social networks in international business. *Journal of International Business Studies*. №. 51 (5). Pp. 714–736.
4. Ahumada-Tello E., Ravina-Ripoll R., Galvez-Albarracin E. G. (2020) Social networks and academic performance self-perception in business sciences students. *Social science*. №. 36 (66). Pp. 105–117.
5. Tiwasing P. (2021) Social media business networks and SME performance: a rural-urban comparative analysis. *Growth and Change*. №. 52 (3). Pp. 1892–1913.

6. Alexandra Ioanid, Cezar Scarlat. (2017) Factors Influencing Social Networks Use for Business: Twitter and YouTube Analysis. *Procedia Engineering*, Volume 181. Pp. 977–983, ISSN 1877-7058. URL: <https://doi.org/10.1016/j.proeng.2017.02.496>.
7. Ilyashenko S. M., Ivanova T. E. (2015) Tools and methods for promoting products on the Internet: an analytical review. *Marketing and innovation management*. No. 3. Pp. 20–32. URL: http://nbuv.gov.ua/UJRN/Mimi_2015_3_4.
8. Shevchuk I. B. (2021) Business at social boundaries: Navch. pos_b. Lviv: Vidavnistvo NNVK «ATB». [in Ukrainian] URL: https://financial.lnu.edu.ua/wp-content/uploads/2021/11/BSM_navchalnyy-posibnyk_2.pdf
9. Varna M. (2023) What is SMM and how does marketing work in social networks. A guide for coworkers. URL: <https://netpeak.net/uk/blog/shcho-take-smm-i-yak-pratsyue-marketing-u-sotsial-nikh-merezhakh-gayd-dlya-pochatktivtsiv/>
10. Marketing in social networks. SMM agency (2022) URL: <https://cases.media/article/marketing-v-socialnikh-merezhakh-smm-agenstvo>
11. Lisa Harris and Alan Rae. (2009) Social networks: the future of marketing for small business. *Journal Of Business Strategyj*. Vol. 30 No. 5.
12. Koirala, Jyoti. (March 27, 2023). Understanding the Use of Cluster Analysis in Business. URL: <https://ssrn.com/abstract=4400674> or <http://dx.doi.org/10.2139/ssrn.4400674>
13. Bezpartochniy M. G. (2014) The use of cluster analysis in the evaluation of the efficiency of trade enterprises. *Trade, commerce, entrepreneurship: collection of scientific works*. Lviv: Lviv Commercial Academy. Issue 17. Pp. 24–27.
14. Jain A. K., Dubes R. C. (1988). *Algorithms for Clustering Data*. Upper Saddle River, NJ: Prentice-Hall, Inc. URL: <http://dl.acm.org/citation.cfm?id=SERIES10022.42779>
15. Jain A.K, Murty M.N., Flynn P.J. (1999) Data Clustering: A Review. *ACM Computing Surveys*. Vol. 31, No. 3. Pp. 264–323. URL: http://www.eecs.northwestern.edu/~yingliu/datamining_papers/survey.pdf
16. Shevchenko S.M., Zhdanova Yu.D., Nehodenko O.V., Kutsuk V.A. (2023) An expert system model for medical screening based on cluster analysis methods. *Moderní aspekty vědy: XXVII. Díl mezinárodní kolektivní monografie*. Mezinárodní Ekonomický Institut s.r.o. Česká republika: Mezinárodní Ekonomický Institut s.r.o. Pp. 478–494. URL: <http://perspectives.pp.ua/public/site/mono/mono-27.pdf>
17. K. A. Abdul Nazeer, M. P. (2009, July 1) Sebastian. Improving the Accuracy and Efficiency of the k-means Clustering Algorithm. *Proceedings of the World Congress on Engineering*. 2009. Vol. I WCE 3. London, U.K.
18. Lotysh O.Ya. (2019) Cluster analysis in industry segmentation. *Bulletin of Odessa National University. Economy*. Issue 5 (78). Volume 24. Pp. 37–42. URL: <https://doi.org/10.32782/2304-0920/5-78-6>
19. Roskladka A.A., Roskladka N.O., Dzigman O.O. (2019) Cluster analysis of the client database of service enterprises. *Agroworld*. No. 16. Pp. 8–17. URL: [https://economics.kntu.kr.ua/pdf/2\(35\)/17.pdf](https://economics.kntu.kr.ua/pdf/2(35)/17.pdf)