

УДК 004.67:004.048:004.89; 005:656.025.2; 656.072; 311
DOI <https://doi.org/10.32782/tnv-tech.2025.1.24>

МЕТОД КЛАСТЕРИЗАЦІЇ НЕЧІТКИХ С-СЕРЕДНІХ В ОРГАНІЗАЦІЇ ПАСАЖИРСЬКИХ ПЕРЕВЕЗЕНЬ У РОЗУМНОМУ МІСТІ

Мацелюх Ю. Р. – аспірант кафедри інформаційних систем та мереж
Національного університету «Львівська політехніка»
ORCID ID: 0000-0002-1721-7703
Scopus-Author ID: 57204941063
Researcher ID: B-6275-2019

Литвин В. В. – доктор технічних наук, професор, завідувач кафедри
інформаційних систем та мереж
Національного університету «Львівська політехніка»
ORCID ID: 0000-0002-9676-0180
Scopus-Author ID: 56446930100
Researcher ID: P-8557-2016

Стрімке зростання обсягів даних, які потребують аналізу, групування і кластеризації, обґрунтовує у світовому масштабі актуальність пошуку методів їх кластеризації, що особливо слушно при організації громадських перевезень. Об'єктом дослідження є процес кластеризації методом нечітких С-середніх великих даних з організації пасажирських перевезень у місті. Предметом – засади оптимізації роботи громадського транспорту та покращення виконання їх графіків. Мета – дослідження особливостей використання методу нечітких С-середніх для оптимізації маршрутів при організації пасажирських перевезень. Завдання: дослідження особливостей метрики методу нечітких С-середніх при організації пасажирських перевезень у розумному місті; проведення поділу на кластери досліджуваного дата-сету; формування алгоритму кластеризації нечітких С-середніх з врахуванням його ефективної метрики.

У результаті досліджень встановлено, що при кластеризації дата-сету про час виконання громадським транспортом кожного перегону в межах кожної години доби для кожного тижня протягом року метод нечітких С-середніх з врахуванням метрики нечіткого розподілу є достатньо простим і ефективним для його обробки, має практичну цінність при виявленні перегонів чи їх груп, де зміни часу виконання кожного перегону для кожного тижня зумовлені змінами трафіку протягом доби або погодними умовами. Завдяки використанню функції належності це сприяє сегментуванню транспортних маршрутів, знижує ймовірність отримання помилок, що підвищує якість прийняття управлінських рішень в організації перевезень.

Висновки. У результаті використання методу нечітких С-середніх з врахуванням метрики нечіткого розподілу при визначенні кількості кластерів було встановлено різні за інтенсивністю трафіку та сезонністю ділянки досліджуваного маршруту. Обґрунтовано, що метод кластеризації нечітких С-середніх достатньо добре враховує специфіку поставленої задачі оптимізації організації пасажирських перевезень у розумному місті, а також є простим і ефективним при обґрунтуванні прийняття управлінських рішень щодо зміни графіків виконання перегонів транспортними засобами.

Ключові слова: пасажирські перевезення; розумне місто; кластерний аналіз; метод нечітких С-середніх; системний аналіз.

Matseliukh Yu. R., Lytvyn V. V. Fuzzy C-means clustering method for organizing passenger transport in a smart city

The rapid growth of data volumes requiring analysis, grouping and clustering justifies the relevance of searching for methods for their clustering on a global scale, which is especially relevant in the organization of public transport. The object of the study is the process of clustering using the fuzzy C-means method of big data in the organization of passenger transport in the city. The subject is the principles of optimizing the operation of public transport and improving the

implementation of their schedules. The goal is to study the features of using the fuzzy C-means method to optimize routes in the organization of passenger transport. Tasks: study the features of the fuzzy C-means method metric in the organization of passenger transport in a smart city; carry out the division of the studied data set into clusters; form a fuzzy C-means clustering algorithm taking into account its effective metric.

As a result of the research, it was found that when clustering a dataset on the time of execution of each public transport run within each hour of the day for each week during the year, the fuzzy C-means method taking into account the fuzzy distribution metric is quite simple and effective for its processing, has practical value in identifying runs or their groups, where changes in the time of execution of each run for each week are due to changes in traffic during the day or weather conditions. Thanks to the use of the membership function, this contributes to the segmentation of transport routes, reduces the probability of errors, which increases the quality of management decision-making in the organization of transportation.

Conclusions. As a result of using the fuzzy C-means method taking into account the fuzzy distribution metric when determining the number of clusters, sections of the studied route were established that differ in traffic intensity and seasonality. It is substantiated that the fuzzy C-means clustering method sufficiently well takes into account the specifics of the task of optimizing the organization of passenger transportation in a smart city, and is also simple and effective in justifying management decisions regarding changes in vehicle travel schedules.

Key words: passenger transport; smart city; clustering analysis; fuzzy C-means method; systems analysis.

Серед методів кластеризації великих даних, що застосовуються в організації пасажирських перевезень у громадському транспорті міст, особливої популярності набуває метод нечітких С-середніх. Його застосовують як для групування споживачів транспортних послуг на основі подібності їх потреб, так і для покращення роботи маршрутів при оптимізації розподілу транспортних засобів. Це сприяє кращому використанню доступних ресурсів, зменшенню викидів, зниженню загальної вартості перевезень. Зростання у містах обсягів даних, які потребують аналізу, групування і кластеризації, обґрунтовує **актуальність** даної **проблеми** на світовому рівні. З потребою в кластеризації даних зустрічаються також при розробці стратегій розумного транспорту, щоб зробити перевезення більш ефективним та орієнтованим на споживачів. Метод нечітких С-середніх відіграє важливу роль як в теорії системного аналізу, так і в розробці інформаційних систем, тому що допомагає приймати оптимальні рішення при моделюванні складних транспортних мереж, покращуючи ефективність роботи з великими даними. Тому у нашій роботі об'єктом досліджень є процес кластеризації методом нечітких С-середніх великих даних з організації пасажирських перевезень місті. Предметом дослідження є засади оптимізації роботи громадського транспорту та покращення виконання їх графіків.

Проаналізувавши **останні дослідження і публікації** у сфері застосування методів та засобів кластерного аналізу [1, 2], бачимо, що у галузі системного аналізу [3–7] багато вчених інтегрують різноманітні методи кластеризації до систем управління розумних транспортом з метою досягнення ефективності, комфортності пасажирських перевезень та зниження вуглецевих викидів [8, 9]. У сучасних дослідженнях відомі науковці в Україні, та за її межами приділяють увагу моделям прийняття рішень, що оптимізують пасажирські перевезення на основі концепції розумного міста та застосовують кластерний аналіз даних для їх організації, а саме: Аша А. [10], Бездек Д. [11], Борейко О. [12], Бублик М. [13, 14], Катренко А. [15], Подлесна Л. [16], Саксена А. [17], Фу Й. [18], Чен Д. [19], Чен З. [20], Целебі М. [21].

Для аналізу наборів даних про пасажирські перевезення переважно використовуються такі методи кластеризації як ієрархічні, розділяючі, на основі щільності,

сітки та штучних нейронних мереж [1–7, 11, 17, 21, 22] для виділення патернів індивідуальних потреб споживачі транспортних послуг, зонування транспортних маршрутів, оптимізації графіків перевезень тощо. Серед методів розділяючої кластеризації вибір припадає на метод К-середніх, який поділяє дані на кластери, знаходячи середню точку в кожному кластері, та метод нечітких С-середніх. Методи розділяючої кластеризації у сфері транспортних перевезень в основному застосовують для дослідження графіків роботи на маршрутах, просторового зонування зупинок на маршрутах з метою їх оптимізації. Також метод нечітких С-середніх часто використовується для побудови гнучких стратегії управління пасажиропотоками при прийнятті рішень щодо врегулювання попиту на транспортні послуги у сучасних містах з динамічною зміною середовища. Особливістю методу нечітких С-середніх є присвоєння елементам набору даних різних ступеней належностей одночасно до всіх кластерів, тобто розраховується ймовірність, з якою кожен елемент належить до кожного з k кластерів. Тому, зважаючи на цю особливість методу, а також на те, що сьогодні все ще невирішеною раніше частиною загальної проблеми є ефективні методи оптимізації транспортних маршрутів, було сформульовано мету роботи: дослідження особливостей використання методу нечітких С-середніх для оптимізації маршрутів при організації пасажирських перевезень громадським транспортом.

У статті вирішуються такі завдання: дослідження особливостей метрик методу нечітких С-середніх при організації пасажирських перевезень у розумному місті; проведення кластеризації набору даних про тривалості виконання перегонів громадським транспортом; формування ефективного алгоритму кластеризації нечітких С-середніх з врахуванням його метрик для обробки наборів даних з громадських перевезень у містах.

Наведемо **результати досліджень та їхнє обговорення**. Провівши дослідження методу нечітких С-середніх та особливостей його метрик, вважаємо, що ключову роль відіграє функція належності елемента до кожного з кластерів, яка визначає ступінь цієї належності у частці від 1 сумарно для всіх кластерів, тобто чим ближче значення до 1, то тим точніше елемент належить до цього кластера. Завдяки цьому метод нечітких С-середніх здатний виділити такі кластери, які неможливо відслідкувати іншими методами, що робить його ефективним при оптимізації маршрутів, покращення їх графіків роботи, тощо. Таким чином, цей метод ефективний у складних системах з великою неоднозначністю та перекриттям між даними, що характерно для даних з організації пасажирських перевезень у громадському транспорті.

Для оцінки якості підбору оптимального значення С-кількості кластерів в методі нечітких С-середніх найчастіше використовується метод розрахунку коефіцієнту нечіткого розподілу (Fuzzy Partition Coefficient, FPC), формула для розрахунку якого наведена в (1):

$$FPC = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C u_{ij}^2, \quad (1)$$

де N – загальна кількість всіх елементів у наборі даних; C – кількість кластерів, які ми хочемо визначити; u_{ij} – ступінь належності i -го елемента до j -го кластера. Саме ступінь належності u_{ij} показує, наскільки сильно i -ий елемент належить до j -го кластера. FPC знаходиться в діапазоні від $1/c$ до 1. Вище значення FPC показує кращу роздільність кластерів, тому вибирається те значення, яке є найбільшим FPC.

Для проведення кластеризації використовувався набір даних про час виконання громадським транспортом кожного перегону, в межах кожної години доби для кожного тижня протягом року. Було виявлено, що дана структура дата-сету, що складається із послідовності: Номер тижня; Геозона; Планований час прибуття; Фактичний час прибуття; Місяць; День; Час; Дата; Тиждень; Година; День тижня; Робочий / вихідний; Перегін 1; Тривалість 1; (...) ; Перегін 10; Тривалість 10, є неоднорідною, складною і потребувала відповідної обробки перед початком аналізу. Тому дані були усереднені в межах кожної доби для кожного тижня протягом року. Отримана і розрахована авторами за зібраними даними структура дата-сету наведена на рис. 1 відповідно у такій послідовності: Номер тижня;; Тривалість перегону 1; ... ; Тривалість перегону 10. З метою пошуку оптимального числа кластерів C для нашого дата-сету із середньодобовим часом виконання кожного перегону для кожного тижня (рис. 1) було використано метод коефіцієнту нечіткого розподілу. Результат розрахунку авторами значення оцінки коефіцієнту нечіткого розподілу FPC наведено на рис. 2, де видно, що максимальне значення $FPC = 0.55$ припадає на число кластерів $k=2$ (Рис. 2), що вказує на оптимальне значення кількості кластерів для кластеризації наших даних.

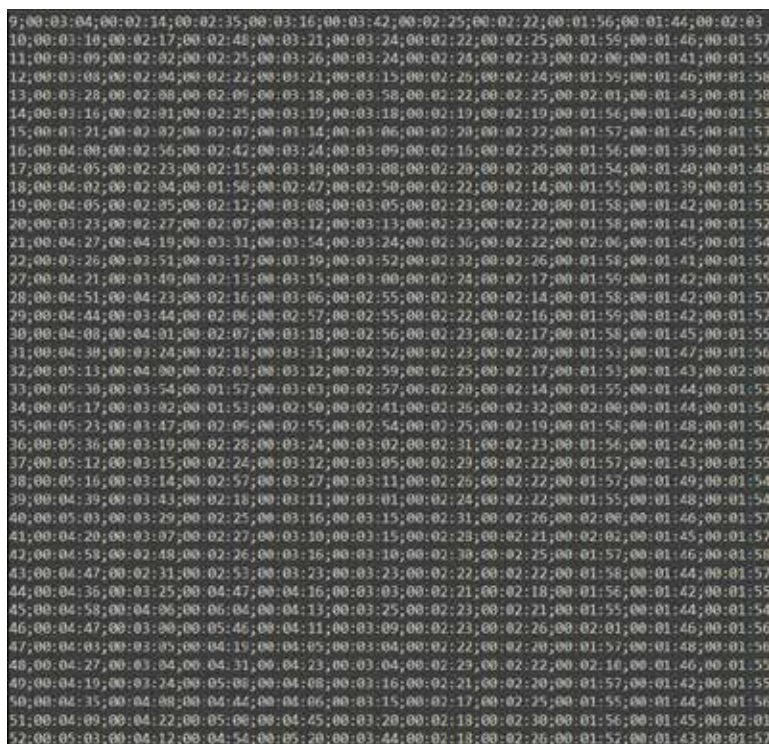


Рис. 1. Середньодобова тривалість виконання кожного перегону транспортними засобами для кожного тижня

Побудований авторами розподіл даних на два кластери ($k=2$) наведено на рис. 3 (відповідно необведена та обведена лінією група точок). Кількість кластерів розраховано за методом коефіцієнту нечіткого розподілу. Перший кластер із тижнів

9–20, 22 і 43 (необведена група точок) утворюють точки зі значеннями середньодобового часу виконання кожного перегону близькими до середнього чи меншими від нього (рис. 1). Другий кластер із рештою тижнів утворює обведена група точок, де час виконання графіків є значно вищими від середньодобових значень (рис. 1) хоча б для одного перегону. При чому, для деяких тижнів (21, 45, 46, 50–52) у кластері спостерігаються перевищення майже у 2 рази середньодобових значень часу виконання графіків на трьох і більше перегонах. Основна маса перевищень часу (затримок) припадає на період року, коли погіршуються погодні умови, а у тижні 21 затримки були зумовлені зростанням трафіку на ділянках, що межують із перегонами у центрі міста, через які проходив об'їзд вулиць із ремонтними роботами.

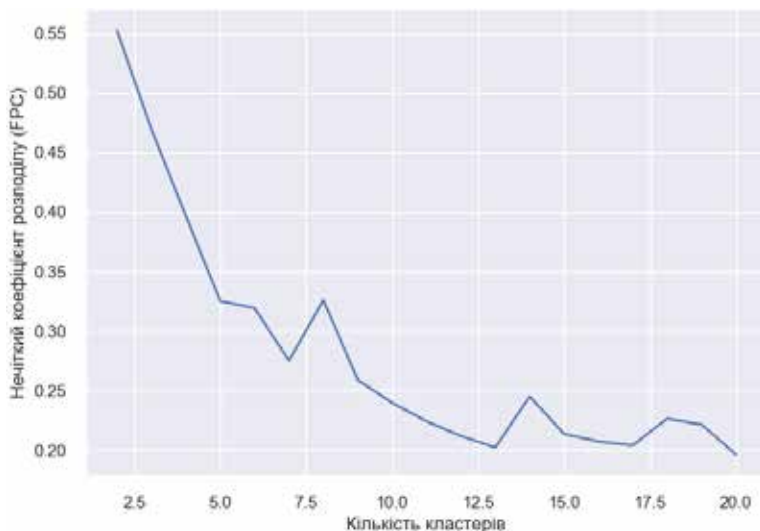


Рис. 2. Залежність коефіцієнту нечіткого розподілу FPC від кількості кластерів k

За допомогою методу кластеризації нечітких С-середніх були виявлені перегони, де протягом тривалого часу на рік (кілька тижнів) відбуваються системні перевищення середньодобових значень часу виконання графіків перегонів порівняно із середньорічними. Це вказує на зростання часу очікування пасажирів на зупинках, що позначається на кількості перевезених пасажирів та на якості наданих послуг, що обґрунтовує рекомендацію прийняття рішення про зміну графіків пасажирських перевезень громадським транспортом у період визначених тижнів з метою його оптимізації.

Отримані результати вказують, що алгоритм кластеризації нечітких С-середніх з врахуванням метрики нечіткого розподілу є достатньо простим і ефективним для обробки даних з громадських перевезень у містах. Він дає змогу більш гнучко та адаптивно виявити ті ділянки маршруту (окремі перегони чи їх групу), де зміни часу виконання кожного перегону транспортними засобами для кожного тижня зумовлені змінами трафіку протягом доби або за різними погодними умовами. Використання методу нечітких С-середніх дає детальнішу картину сегментування транспортних маршрутів, знижує ймовірність отримання помилки при аналізі різних ділянок транспортної мережі завдяки наявності функції належності. Все це підвищує якість прийняття стратегічних рішень в організації транспортних перевезень.

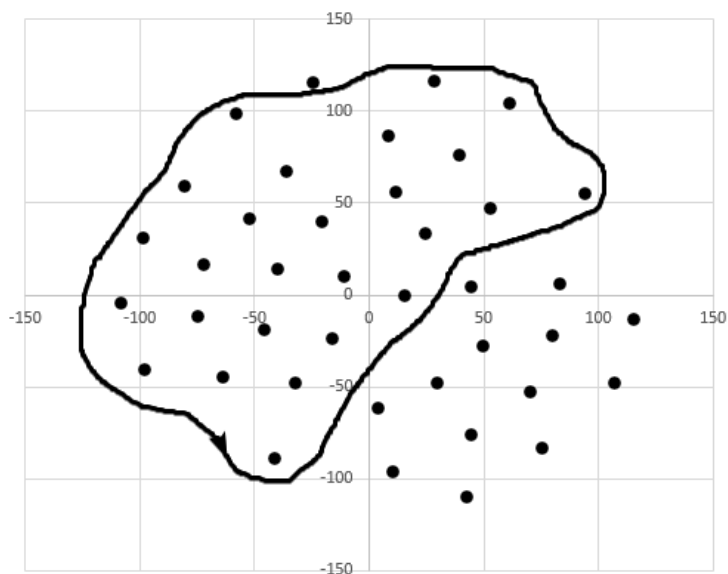


Рис. 3. Поділ на 2 кластери за методом нечітких С-середніх середньодобової тривалості виконання кожного перегону кожного тижня

Отже, можна зробити **висновок**, що у роботі з метою виявлення особливостей використання методу нечітких С-середніх для оптимізації маршрутів при організації пасажирських перевезень громадським транспортом було проведено кластерний аналіз дата-сету про час виконання перегонів громадським транспортом у середньорозмірному місті. У результаті цього було виявлено окремі ділянки (перегони, зони) маршруту з різною інтенсивністю, на які високий вплив мають трафік, а також сезонність. Це дало змогу встановити, що метод кластеризації нечітких С-середніх достатньо добре враховує специфіку поставленої задачі оптимізації організації пасажирських перевезень у розумному місті, а також є простим і ефективним при обґрунтуванні прийняття управлінських рішень щодо зміни графіків виконання перегонів транспортними засобами.

Перспективою подальших досліджень є дослідження методу нечітких С-середніх для виявлення неефективних маршрутів з низьким завантаженням окремих ділянок транспортної мережі чи їх перевантаженням.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ:

1. Ezugwu A. E., Ikotun A. M., Oyelade O. O., Abualigah L., Agushaka J. O., Eke C. I., Akinyelu A. A. A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. *Engineering Applications of Artificial Intelligence*. 2022. № 110, 104743. DOI: 10.1016/j.engappai.2022.104743
2. Singh J., Singh D. A comprehensive review of clustering techniques in artificial intelligence for knowledge discovery: Taxonomy, challenges, applications and future prospects. *Advanced Engineering Informatics*. 2024. № 62, 102799. DOI: 10.1016/j.aei.2024.102799
3. Bublyk M., Kowalska-Styczeń A., Lytvyn V., Vysotska V. The Ukrainian economy transformation into the circular based on fuzzy-logic cluster analysis. *Energies*. 2021. № 14(18). P. 5951. DOI: 10.3390/en14185951

4. Ester M., Kriegl H.-P., Sander J., Xu X. *A density-based algorithm for discovering clusters in large spatial databases with noise*. Proc. of the 2-nd Int. Conf. on Knowledge Discovery and Data Mining. 1996. P. 226–231.
5. Kowalska-Styczeń A., Bublyk M., Lytvyn V. Green innovative economy remodeling based on economic complexity. *Journal of Open Innovation: Technology, Market, and Complexity*. 2023. V.9 (3), 100091. DOI: 10.1016/j.joitmc.2023.100091
6. Jain A. K. Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*. 2010. № 31(8). P. 651–666. DOI: 10.1016/j.patrec.2009.09.011
7. Kohonen T. *Self-Organizing Maps* (3 ed). Bon, Springer, 2001. 554 p.
8. Matseliukh Y., Bublyk M., Bosak A., Naychuk-Khrushch M. The role of public transport network optimization in reducing carbon emissions. *CEUR Workshop Proceedings*. 2024. Vol. 3723. P. 340–364. URL: <https://ceur-ws.org/Vol-3723/paper19.pdf>
9. Koshtura D., Bublyk M., Matseliukh Y., Dosyn D., Chyrun L., Lozynska O., Karpov I., Peleshchak I., Maslak M., Sachenko O. Analysis of the demand for bicycle use in a smart city based on machine learning. *CEUR workshop proceedings*. 2020. Vol. 2631, P. 172–183. URL: <https://ceur-ws.org/Vol-2631/paper13.pdf> (дата звернення: 01.02.2025).
10. Asha A., Arunachalam R., Poonguzhali I., Urooj S., Alelyani S. Optimized RNN-based performance prediction of IoT and WSN-oriented smart city application using improved honey badger algorithm. *Measurement*. 2023. № 210. P. 112505. DOI: <https://doi.org/10.1016/j.measurement.2023.112505>
11. Bezdek J. C. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Bon, Springer, 1981, 245 p.
12. Boreiko O., Teslyuk V. Structural model of passenger counting and public transport tracking system of smart city. *Perspective Technologies and Methods in MEMS Design, Proc. of Int. Conf.* 2016. P. 124–126. DOI: <https://doi.org/10.1109/MEMSTECH.2016.7507533>
13. Bublyk M., Udovychenko T., Medvid R. Concept of smart specialization in the context of the development of Ukraine's economy. *Economics. Ecology. Socium*. 2019. № 3(2). P. 55–61. DOI: 10.31520/2616-7107/2019.3.2-6
14. Bublyk M.I., Dulyaba N.I., Petryshyn N.Y., Drymalovska K.V. Analysis of the development of educational and scientific activity of institutions of higher education. *Economic analysis*. 2018. № 28 (1). P. 30–39.
15. Katrenko A., Krislata I., Veres O., Oborska O., Basyuk T., Vasyliuk A., Rishnyak I., Demyanovskyi N., Meh O. Development of traffic flows and smart parking system for smart city. *CEUR Workshop Proceedings*. 2020. Vol. 2604. P. 730–745. URL: <https://ceur-ws.org/Vol-2604/paper50.pdf>
16. Podlesna L., Bublyk M., Grybyk I., Matseliukh Y., Burov Y., Kravets P., Lozynska O., Karpov I., Peleshchak I., Peleshchak R. Optimization model of the buses number on the route based on queueing theory in a Smart City. *CEUR Workshop Proceedings*. 2020. Vol. 2631. P. 502–515. URL: <https://ceur-ws.org/Vol-2631/paper37.pdf> (дата звернення: 01.02.2025).
17. Saxena A., Prasad M., Gupta A., Bharill N., Patel O. P., Tiwari A., Er M. J., Ding W., Lin C. A review of clustering techniques and developments. *Neurocomputing*. 2017. № 267. P. 664–681. DOI: 10.1016/j.neucom.2017.06.053
18. Fu Y., Wang X., Meng F., Wang S., Song Y., Wang Y. Behavioural loyalty analysis of bus passengers using multi-source data fusion. *Journal of Transport Geography*. 2025. № 123. P. 104143. <https://doi.org/10.1016/j.jtrangeo.2025.104143>
19. Chen J., Zheng L., Hu Y., Wang W., Zhang H., Hu X. Traffic flow matrix-based graph neural network with attention mechanism for traffic flow prediction. *Inf. Fusion*. 2024. № 104. P. 102146. <https://doi.org/10.1016/j.inffus.2023.102146>
20. Chen Z., Wen J., Geng Y. Predicting future traffic using hidden Markov models. *Proc., 2016 IEEE 24th Int. Conf. on Network Protocols (ICNP)*. 2016. P. 11.

21. Celebi M. E., Kingravi H. A., Vela P. A. A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Systems With Applications*. 2012. № 40(1). P. 200–210. DOI: 10.1016/j.eswa.2012.07.021

22. Chukhray N. Shakhovska N., Mrykhina O., Bublyk M., Lisovska L. Consumer aspects in assessing the suitability of technologies for the transfer. *Computer sciences and information technologies, CSIT 2019: Proc. of Int. Sci. Tech. Conf.* 2019. V. 3. P. 142–147. DOI: 10.1109/STC-CSIT.2019.8929879

REFERENCES:

1. Ezugwu, A. E., Ikotun, A. M., Oyelade, O. O., Abualigah, L., Agushaka, J. O., Eke, C. I., & Akinyelu, A. A. (2022). A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. *Engineering Applications of Artificial Intelligence*, 110, 104743. DOI: 10.1016/j.engappai.2022.104743

2. Singh, J., Singh, D. (2024). A comprehensive review of clustering techniques in artificial intelligence for knowledge discovery: Taxonomy, challenges, applications and future prospects. *Advanced Engineering Informatics*, 62, 102799. DOI: 10.1016/j.aei.2024.102799

3. Bublyk, M., Kowalska-Styczeń, A., Lytvyn, V., Vysotska, V. (2021). The Ukrainian economy transformation into the circular based on fuzzy-logic cluster analysis. *Energies*, 14 (18), 5951. DOI: 10.3390/en14185951

4. Ester, M., Kriegel, H.-P., Sander, J., Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Proc. of the 2nd Int. Conf. on Knowledge Discovery and Data Mining*, 226–231.

5. Kowalska-Styczeń, A., Bublyk, M., Lytvyn, V. (2023). Green innovative economy remodeling based on economic complexity. *Journal of Open Innovation: Technology, Market, and Complexity*, 9(3), 100091. DOI: 10.1016/j.joitmc.2023.100091

6. Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31 (8), 651–666. DOI: 10.1016/j.patrec.2009.09.011

7. Kohonen, T. (2001). *Self-Organizing Maps*, 3 ed., Bon, Springer.

8. Matseliukh, Y., Bublyk, M., Bosak, A., Naychuk-Khrushch, M. (2024). The role of public transport network optimization in reducing carbon emissions. *CEUR Workshop Proceedings*, Vol. -3723, 340–364. Retrieved from <https://ceur-ws.org/Vol-3723/paper19.pdf>.

9. Koshtura, D., Bublyk, M., Matseliukh, Y., Dosyn, D., Chyrun, L., Lozynska, O., Karpov, I., Peleshchak, I., Maslak, M., Sachenko, O. (2020). Analysis of the demand for bicycle use in a smart city based on machine learning. *CEUR Workshop Proceedings*, Vol. -2631, 172–183. Retrieved from <https://ceur-ws.org/Vol-2631/paper13.pdf>.

10. Asha, A., Arunachalam, R., Poonguzhali, I., Urooj, S., Alelyani, S. (2023). Optimized RNN-based performance prediction of IoT and WSN-oriented smart city application using improved honey badger algorithm. *Measurement*, 210, 112505. DOI: 10.1016/j.measurement.2023.112505

11. Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*, Bon, Springer.

12. Boreiko, O., Teslyuk, V. (2016). Structural model of passenger counting and public transport tracking system of smart city. *Perspective Technologies and Methods in MEMS Design, Proc. of Int. Conf.*, 124–126, DOI: 10.1109/MEMSTECH.2016.7507533

13. Bublyk, M., Udovychenko, T., Medvid, R. (2019). Concept of smart specialization in the context of the development of Ukraines economy. *Economics. Ecology. Socium*, 3(2), 55–61. DOI: 10.31520/2616-7107/2019.3.2-6

14. Bublyk, M., Dulyaba, N., Petryshyn, N., Drymalovska, K. (2018). Analysis of the development of educational and scientific activity of institutions of higher education. *Economic analysis*, 28(1), 30–39.

15. Katrenko, A., Krislata, I. Veres, O., Oborska, O., Basyuk, T., Vasyliuk, A., Rishnyak, I., Demyanovskyi, N., Meh, O. (2020). Development of traffic flows and smart parking system for smart city. *CEUR Workshop Proceedings*, Vol. -2604, P. 730–745. Retrieved from <https://ceur-ws.org/Vol-2604/paper50.pdf>.
 16. Podlesna, L., Bublyk, M., Grybyk, I., Matseliukh, Y., Burov, Y., Kravets, P., Lozynska, O., Karpov, I., Peleshchak, I., Peleshchak, R. (2020). Optimization model of the buses number on the route based on queueing theory in a Smart City. *CEUR Workshop Proceedings*, Vol. -2631, 502–515. Retrieved from <https://ceur-ws.org/Vol-2631/paper37.pdf>.
 17. Saxena, A., Prasad, M., Gupta, A., Bharill, N., Patel, O. P., Tiwari, A., Er, M. J., Ding, W., Lin, C. (2017). A review of clustering techniques and developments. *Neurocomputing*, 267, 664–681. DOI: 10.1016/j.neucom.2017.06.05
 18. Fu, Y., Wang, X., Meng, F., Wang, S., Song, Y., & Wang, Y. (2025). Behavioural loyalty analysis of bus passengers using multi-source data fusion. *Journal of Transport Geography*, 123, 104143. DOI:10.1016/j.jtrangeo.2025.10414
 19. Chen, J., L. Zheng, Y. Hu, W. Wang, H. Zhang, and X. Hu. (2024). Traffic flow matrix-based graph neural network with attention mechanism for traffic flow prediction. *Inf. Fusion*, 104, 102146. DOI: 10.1016/j.inffus.2023. 102146.
 20. Chen, Z., J. Wen, and Y. Geng. (2016). Predicting future traffic using hidden Markov models. *Proc., 2016 IEEE 24th Int. Conf. on Network Protocols*, 11
 21. Celebi, M. E., Kingravi, H. A., Vela, P. A. (2012). A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Systems With Applications*, 40(1), 200–210. DOI: 10.1016/j.eswa.2012.07.021
 22. Chukhray, N. Shakhovska, N., Mrykhina, O., Bublyk, M., Lisovska L. (2019). Consumer aspects in assessing the suitability of technologies for the transfer. *Computer sciences and information technologies, CSIT 2019: Proc. of Int. Sci. Tech. Conf.*, 3, 142–147. DOI: 10.1109/STC-CSIT.2019.8929879
-