**УДК 004.428.2**

*Лухверчик С. А., здобувач 2 курсу*
*ОС «Магістр» спеціальності*
*F3 Комп'ютерні науки,*
*Потапова Н. А., канд. екон. наук, доцент,*
*доцент кафедри інформаційних технологій*

# АНАЛІТИЧНІ МЕТОДИ ОБРОБКИ ВЕЛИКИХ ДАНИХ У ЛОГІСТИЧНИХ СИСТЕМАХ

*Донецький національний університет імені Василя Стуса, м. Вінниця*

Цифрова трансформація логістичних процесів призводить до стрімкого зростання обсягів інформації, яку генерують транспортні засоби, складські системи, модулі відстеження замовлень та аналітичні сервіси. Дані надходять у вигляді координат, часових міток, атрибутів замовлень і телеметричних параметрів, формуючи різнорідні інформаційні потоки.

Для логістичної галузі характерні проблеми неповноти, розрізненості та нерівномірного просторового розподілу даних, що ускладнює розробку ефективних моделей підтримки рішень. Саме тому методи Big Data стають ключовим інструментом для виявлення закономірностей, прогнозування попиту та формування структурованих логістичних зон.

На основі аналізу реальних логістичних наборів даних виконано структуризацію просторової інформації та оцінено вплив попередньої підготовки даних на якість подальшого аналізу. Виявлено, що пропуски, дублікати координат і аномальні значення значно спотворюють результати моделювання, тому критичними є етапи очищення, нормалізації та фільтрації.

Ключовим напрямом дослідження стала кластеризація точок доставки – один з найпоширеніших методів просторової аналітики в логістиці. Алгоритм K-Means продемонстрував ефективність у випадках відносно рівномірного розподілу точок, даючи змогу формувати компактні й інтерпретовані кластери. Водночас DBSCAN виявив здатність коректно групувати точки в умовах нерівномірної щільності, що є типовою проблемою для кур'єрських та регіональних служб доставки. Результати показали, що кластеризація допомагає:

1) формувати логістичні мікрорегіони для подальшої маршрутизації;
2) оцінювати динаміку та щільність попиту;
3) візуалізувати «гарячі зони» доставки;
4) визначати території з потенційними логістичними ризиками;
5) знижувати навантаження на транспортні ресурси завдяки правильному групуванню точок.

У межах дослідження було побудовано візуалізації кластерних моделей, що дають змогу наочно представити логістичну структуру території.

На рис. 1 наведено приклади типових результатів кластеризації у логістиці: компактні центроїдні кластери (K-Means) та кластери довільної форми (DBSCAN), що відображають просторову нерівномірність.

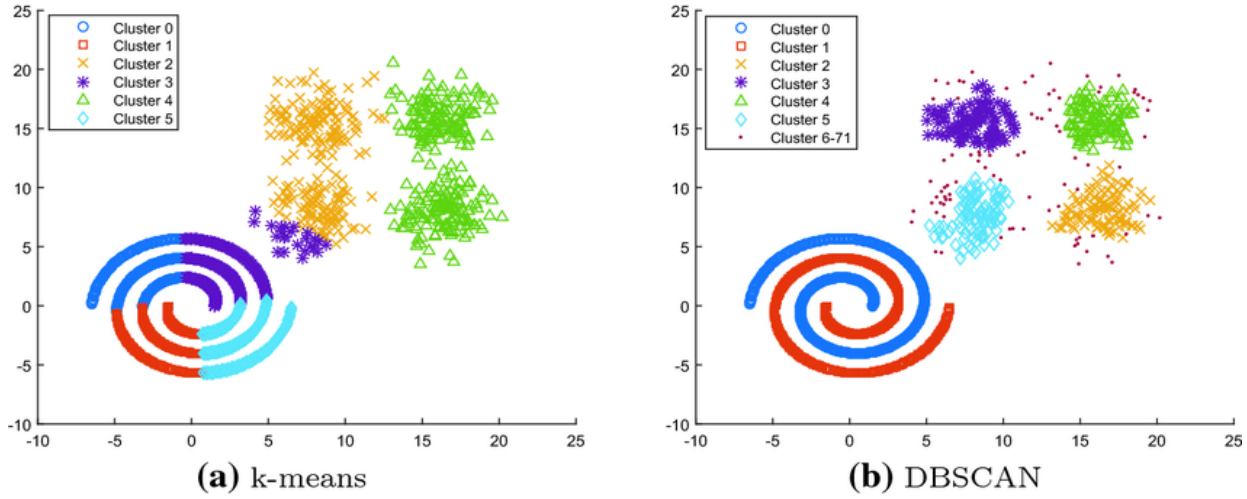



*Рисунок 1 – Порівняння результатів кластеризації методом K-Means та DBSCAN*

Аналітичний підхід до обробки Big Data у логістиці забезпечує підвищення точності оцінювання попиту, прозорість логістичних операцій та можливість формувати адаптивні регіональні моделі, придатні для практичного застосування в кур'єрських службах, дистрибуційних компаніях і системах міської доставки.

Методи аналітики великих даних відіграють важливу роль у формуванні сучасних логістичних рішень. Проведене дослідження показало, що кластеризація є одним з найефективніших інструментів попереднього структурування даних, яке дає змогу суттєво покращити подальші процеси моделювання та управління. Використання K-Means та DBSCAN забезпечує можливість адаптивного поділу логістичного простору залежно від щільності та характеристик розташування точок доставки.

Отже, Big Data-підходи створюють основу для розробки більш гнучких, стійких і прозорих логістичних систем, що відповідають викликам сучасної цифрової економіки.

## Список використаних джерел

1. Han J., Kamber M., Pei J. Data Mining: Concepts and Techniques. Amsterdam: Elsevier, 2012. 740 p.

2. Provost F., Fawcett T. Data Science for Business. Sebastopol: O'Reilly Media, 2013. 414 p.

3. Gandomi A., Haider M. Beyond the hype: Big Data concepts, methods, and analytics. *International Journal of Information Management.* 2015. URL: https://doi.org/10.1016/j.ijinfomgt.2014.10.007 (дата звернення: 22.11.2025).

4. Raghupathi W., Raghupathi V. Big Data Analytics in Healthcare: Promise and Potential. *Health Information Science and Systems.* 2014. № 3. URL: https://doi.org/10.1186/2047-2501-2-3 (дата звернення: 22.11.2025).

5. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise (DBSCAN) / M. Ester, H.-P. Kriegel, J. Sander, X. Xu. *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96).* Portland, USA, 1996. P. 226–231. URL: https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf (дата звернення: 22.11.2025).