

МЕТОДИ ВИКОРИСТАННЯ ТА ПРИШВИДШЕННЯ ОБРОБКИ ВЕЛИКИХ МАСИВІВ ДАНИХ

Кануков Д. С., Жемір О. В., Дорошев Я. О., Партика С. О.
Харківський національний університет радіоелектроніки, Харків, Україна

Кожен день генеруються великі об'єми даних, через різноманітність їх структури та характеристик ускладнюється їх обробка і класифікація. Через саму природу Big Data, як правило, їх важко обробляти, керувати ними та ефективно використовувати. Середовища Big Data, як правило, складні, з кількома застосовуваними системами та інструментами, які потрібно добре налаштувати для сумісної праці. Самі такі данні також мають складну структуру, особливо коли вони мають великий розмір й різноманітні, або включають поточкові данні. Проте основною перевагою Big Data є статистична потужність, оскільки ці данні мають велику кількість характеристик, що покращує їх кореляцію. Отримання закономірностей дозволяє виявляти нові тенденції для ведення бізнесу, виявляти вплив людини на екологію та розробляти різноманітні алгоритми управління.

Метою доповіді є демонстрування пришвидшення обробки даних за допомогою використання інтелектуальної системи прийняття рішень у порівнянні з традиційним програмним забезпеченням для обробки даних.

В доповіді наведено використані ресурси для порівняння різноманітних методів обробки з однаковою кількістю та складністю масивів даних. Оскільки кожен з методів обробки даних потребує свій рівень кваліфікації, рівень потужності обчислюваної техніки та потребує багато часу, то можна зробити висновок про актуальність розробки нових методів роботи з Big Data та застосованих для цього технологій.

Список літератури

1. <https://www.techtarget.com/searchdatamanagement/The-ultimate-guide-to-big-data-for-businesses>
2. Seth Stephens-Davidowitz. Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are. 2017. <http://www.lib.uni-corvinus.hu/eng/content/everybody-lies-big-data-new-data-and-what-internet-can-tell-us-about-who-we-really-are>.
3. Rob Kitchin, What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets DOI: <https://doi.org/10.1177/2053951716631130>.
4. Priyesh Vijayan, Yash Chandak, Mitesh M. Khapra, Srinivasan Parthasarathy and Balaraman Ravindran. Scaling Graph Propagation Kernels for Predictive Learning DOI: <https://doi.org/10.3389/fdata.2022.616617>.
5. Édition en Anglais, Big Data: Principles and Best Practices of Scalable Real Time Data Systems. 2015 <https://www.audible.fr/pd/Big-Data-Principles-and-Best-Practices-of-Scalable-Realtime-Data-Systems-Livre-Audio/B07K337CRT>