

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В ОБРОБЦІ ДАНИХ

Кузьменко Г.О., Штангей С.В.

Кафедра інфокомунікаційної інженерії ім. В.В. Поповського,
Харківський національний університет радіоелектроніки,
Україна.

E-mail: heorhii.kuzmenko@nure.ua
svitlana.shtanhei@nure.ua

Abstract

The abstract summarizes an exploration into data processing's pivotal role in today's context, highlighting its fundamental concepts, analysis, and the significance of data collection and utilization. It delves into innovative technologies like artificial intelligence (AI) and machine learning, elucidating their roles in data processing. The discussion extends to various machine learning models and their practical applications in data analysis. Additionally, it defines Big Data and emphasizes the need for efficient processing of substantial data volumes, spotlighting tools such as Hadoop and Mapreduce for this purpose. The annotation explores the Internet of Things (IoT) and its role in data collection, outlining associated challenges and future prospects. Real-world applications of these innovations in business analytics, decision-making, medical research, and data processing are detailed. Furthermore, it emphasizes the criticality of data security in processing and storage. Finally, the abstract concludes with a summary of key findings and future forecasts for the advancement of innovative data processing technologies.

Вступ

1. Визначення та важливість обробки даних.

Процеси, що забезпечують роботу інформаційної системи будь-якого призначення, складаються з таких основних етапів:

- введення інформації із зовнішніх або внутрішніх джерел;
- обробка вхідної інформації й структуризація її в зручному виді;
- вивід інформації для передачі споживачам або в іншу систему;
- повторне введення або корекція незадовільно зібраної вхідної інформації. [2]

Обробка даних – це впорядкований процес перетворення даних у відповідності з алгоритмом розв'язання завдання. Процес обробки даних полягає в отриманні одних «інформаційних об'єктів» з інших «інформаційних об'єктів» шляхом виконання деяких алгоритмів і є однією з основних операцій, здійснюваних над даними. [1]

З точки зору реалізації на основі сучасних досягнень обчислювальної техніки виділяють наступні види обробки даних:

- послідовна обробка, застосовувана в традиційній архітектурі комп'ютера з одним процесором;
- паралельна обробка, застосовувана при наявності декількох процесорів в комп'ютері;
- конвеєрна обробка, пов'язана з використанням в архітектурі комп'ютера одних і тих же ресурсів для розв'язання різних завдань, причому якщо ці завдання тотожні, то це послідовний конвеєр, якщо завдання однакові – векторний конвеєр.

Реалізація всіх дій, які виконуються в процесі обробки даних, здійснюється за допомогою різноманітних програмних засобів. [1]

Після розв'язання задач обробки даних результат повинен бути виданий кінцевим користувачам в зручній для них формі. Ця операція реалізується в процесі видачі (використання) даних. Видача (використання) даних, як правило, здійснюється за допомогою технічних пристроїв у вигляді текстів, таблиць, графіків, зображень і т. ін. [1]

Огляд інноваційних технологій

1. Штучний інтелект і машинне навчання.

Однією з передових інформаційних технологій сучасності є штучний інтелект. Як робоче визначення науковці пропонують таке: **штучний інтелект** – це один із напрямів інформатики, метою якого є розробка апаратно-програмних засобів, які дозволяють користувачеві непрограмістові ставити й вирішувати свої, що традиційно вважаються інтелектуальними, завдання, спілкуючись з електронно-обчислювальною машиною на обмеженій підмножині природної мови. [3]

Штучний інтелект має велике значення в обробці даних, оскільки надає можливість автоматизувати аналіз, розпізнавання шаблонів та розуміння великих обсягів інформації. Його потужності допомагають покращити прогностичні моделі, забезпечують персоналізовану рекомендацію та адаптацію у багатьох сферах, у тому числі в бізнес і медицину. Використання штучного інтелекту сприяє інноваціям та прогресу. [4]

2. Big data та інструменти для його обробки.

Великі дані (Big Data) – позначення структурованих и неструктурованих даних величезних обсягів і значного розмаїття, що піддаються ефективній обробці програмних інструментів, які горизонтально масштабуються. [5]

У сукупність підходів і технологій включають засоби масової паралельної обробки невизначено-структурованих даних, серед них є популярними NoSQL, алгоритми MapReduce і засоби проекту Hadoop. **MapReduce** — модель розподілених обчислень у комп'ютерних кластерах, представлена компанією Google. Згідно з цією моделлю, додаток розділяється на значну кількість однакових елементарних завдань(рис. 1), що виконуються на вузлах кластера і потім, природнім шляхом зводяться у кінцевий результат. [5]

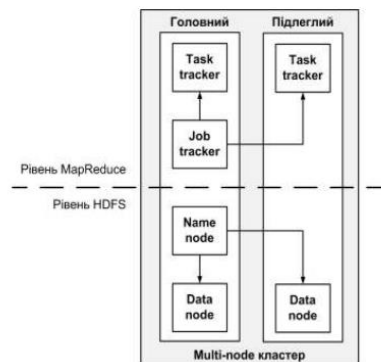


Рис. 1. Обчислювальна парадигма MapReduce

Hadoop — проект фонду Apache Software Foundation, набір утилітів, бібліотек і фреймворків, що вільно розповсюджується, для розробки і виконання розподілених програм, які працюють на кластерах із сотень і тисяч вузлів. Вважається однією з основоположних технологій більшості даних. [5]

3. ІОТ та збір даних.

Інтернет речей (Internet of Things, скорочено IoT) — це глобальна мережа підключених до Інтернету речей — пристроїв, оснащених сенсорами, датчиками, засобами передавання сигналів. Інтернет речей об'єднує реальні речі в віртуальні системи, здатні вирішувати абсолютно різні завдання. Ключова ідея — з'єднати між собою всі об'єкти, які можна з'єднати, підключити їх до мережі для збирання даних і прийняття рішень на їх основі. [6]

Інтернет речей може викликати величезні зміни у повсякденному житті, надавши звичайним користувачам абсолютно новий рівень комфорту. Але якщо елементи такої системи не будуть належним чином захищені від несанкціонованого втручання, за допомогою надійного

криптографічного алгоритму, замість користі вони принесуть шкоду, надавши кіберзлочинцям лазівку для підриву інформаційної безпеки. [7]

Застосування в реальних сферах

1. Обробка даних у бізнесі.

Необхідною умовою сучасного статистичного аналізу даних як основного процесу бізнес-аналітики є ефективне використання комп'ютерних програм, від функціональної повноти й алгоритмічної продуманості яких залежить підсумкова інтерпретація результатів дослідження та надійність висновків. Однією з найзручніших для досягнення цієї мети є вільно розповсюджене статистичне середовище R, яке є найбільш повним, надійним, динамічно розвивається та об'єднує мову програмування високого рівня і потужні бібліотеки програмних модулів для обчислювальної та графічної обробки даних. [8]

R – це універсальна мова програмування, розроблена для застосування в галузях сучасної бізнес-аналітики з аналізом даних, класичної статистики і якісної модерної графіки. Завдяки своїй великій бібліотеці пакетів, що безперервно розширюється, мова підтримує широкий спектр статистичних та чисельних методів. [8]

2. Медицина та наука.

На практиці для лікаря на сьогоднішній день обробка й аналіз даних зводиться до вирішення наступних завдань:

- мати уявлення про основні статистичні методи;
- освоїти відповідний пакет, що буде використовуватися для аналізу й обробки даних;
- провести власне цей аналіз та вміти правильно інтерпретувати отримані результати.

Завдяки використанню комп'ютерів і широкому впровадженню сучасних комп'ютерних технологій докорінно змінився процес обробки й аналізу медичних даних. Застосування комп'ютерної техніки робить достатньо складні методи аналізу медичних даних більш доступними і наочними. Вже не потрібно вручну виконувати трудомісткі розрахунки, будувати таблиці і графіки - всю цю роботу взяв на себе комп'ютер, а людині залишилася лише творча робота: постановка задач, вибір методів їх вирішення та інтерпретація результатів.

3. Кібербезпека та захист даних

Безпека системи є головною складовою зменшення кількості злочинів, здійснених у кіберпросторі. [9] Тому, забезпечення безпеки при обробці та збереженні даних має критичне значення в сучасному цифровому світі. Це не лише захист від кібератак, а й збереження конфіденційності, цілісності та доступності інформації. Інформація є ключовим активом для бізнесу, уряду та приватних осіб, тому захист від несанкціонованого доступу, втрати чи порушень є вирішальними для запобігання втратам довіри, фінансових втрат та порушень приватності. [10]

Ефективна кібербезпека вимагає постійного вдосконалення стратегій, використання передових технологій шифрування, автентифікації та моніторингу, а також розуміння потенційних загроз для запобігання проблемам, які можуть виникнути. [10]

Висновки

Під час опрацювання теми “Інформаційні технології в обробці даних”, ми визначили, що обробка даних, заснована на інноваційних технологіях, є фундаментальним елементом сучасного світу, де збір, аналіз та використання інформації має перевагу у вирішенні різних завдань.

Розглянуті технології, такі як штучний інтелект, машинне навчання, Big Data та Інтернет речей, мають великий потенціал у різних сферах. ШІ та машинне навчання дозволяють розвивати нові підходи до обробки даних, тоді як Big Data вимагає ефективних інструментів для обробки великих обсягів інформації. IoT створює можливість збирати дані в реальному часі та впливати на велику кількість сфер життя.

Такий широкий спектр застосування інноваційних технологій вимагає відповідної уваги до захисту даних. Забезпечення безпеки при обробці та зберіганні даних є критичним для запобігання кіберзагрозам та збереження довіри, цілісності та конфіденційності інформації. Виходячи з цього, інноваційні технології потребують не тільки постійного розвитку, а й удосконалення заходів безпеки для забезпечення стабільності та надійності.

Я гадаю, що у майбутньому одним з ключових аспектів буде постійне збільшення швидкості та ефективності обробки даних, що дозволить використовувати їх у реальному часі для прийняття рішень та вирішення складних завдань. Розширення сфери застосування штучного інтелекту принесе інновації у багатьох галузях, починаючи від медицини до транспорту, відкриваючи нові можливості та вирішуючи складні завдання.

Крім того, майбутнє має бути спрямоване на розвиток захисту приватності. Зростання уваги до конфіденційності може призвести до створення строгих стандартів та регулювань у цій сфері, забезпечуючи більшу відповідальність на захист інформації. Такі рухи в майбутньому сприятимуть створенню більш безпечного та етичного середовища для обробки даних, що є головним для підтримки стабільності.

Література

1. Бонч-Бруєвич Г.Ф., Носенко Т.І. Організація та обробка електронної інформації: навчальний посібник./Бонч-Бруєвич Г.Ф., Носенко Т.І.; Київ. ун-т ім. Б.Грінченка, ін-т суспільства, каф. інформатики. – К.: [Київ. ун-т ім. Б. Грінченка], 2013. – 108 с.
2. Грицунов О. В. Інформаційні системи та технології: навч. посіб. для студентів за напрямом підготовки «Транспортні технології» / О. В. Грицунов; Харк. нац. акад. міськ. госп-ва. – Х.: ХНАМГ, 2010. – 222 с. ISBN 978-966-695-195-6
3. Kryvytskyi Yu. – Ph.D in Law, Associate Professor, Associate Professor of the Department of Theory of State and Law of the National Academy of Internal Affairs, Kyiv, Ukraine ORCID: <https://orcid.org/0000-0002-7063-4725>
4. Методи та системи штучного інтелекту: Навчальний посібник для студентів напряму підготовки 6.050101 «Комп'ютерні науки» / Уклад. : А.С. Савченко, О. О. Синельников. – К. : НАУ, 2017. – 190 с.
5. ЛНТУ. Технології обробки великих даних (Big Data). Тема 16. 2010. - 24 с. URL: https://e-tk.lntu.edu.ua/pluginfile.php/20333/mod_resource/content/0/Тема%2016.%20Технології%20обробки%20великих%20даних%20%28Big%20Data%29.pdf
6. Технології інтернету речей. Навчальний посібник [Електронний ресурс]: навч. посіб. для студ. спеціальності 126 «Інформаційні системи та технології», спеціалізація «Інформаційне забезпечення робототехнічних систем» / Б. Ю. Жураковський, І.О. Зенів; КПІ ім. Ігоря Сікорського. – Електронні текстові дані (1 файл: 12,5 Мбайт). – Київ: КПІ ім. Ігоря Сікорського, 2021. – 271 с.
7. Державний університет "Житомирська політехніка". Вступ до Інтернету речей. ЛЕКЦІЯ з навчальної дисципліни «Архітектура і технології IoT». № 1 - 13 с. URL: https://learn.ztu.edu.ua/pluginfile.php/68838/mod_resource/content/2/Л1-1.pdf
8. Сидорова А. В., Біленко Д. В., Буркіна Н. В. Бізнес-аналітика: навчально-методичний посібник. Вінниця: ДонНУ імені Василя Стуса. 2019. 104 с.
9. Яремчук Ю. Є. Комплексні системи захисту інформації: навчальний посібник / Яремчук Ю. Є., Павловський П. В., Катаєв В. С., Сінюгін В. В. – Вінниця: ВНТУ, 2017. – 120 с.
10. Методологія захисту інформації. Аспекти кібербезпеки: підручник/ Г.М. Гулак – К.: Видавництво НА СБ України, 2020. – 256 с.
11. О.В. Гойко. СУЧАСНІ ТЕХНОЛОГІЇ ОБРОБКИ Й АНАЛІЗУ МЕДИЧНИХ ДАНИХ: стаття./ Медична інформатика та інженерія, № 4, 2009 - 6 ст. URL: <https://core.ac.uk/download/pdf/276621005.pdf>