

Ярошенко О.І. к.е.н., доцент
Чернівецький національний університет імені Юрія Федьковича, м. Чернівці
Мелашенко А.О., к.ф.-м.н., старший наук. співр.
Інститут кібернетики імені В.М. Глушкова НАН України, м. Київ

ВИКОРИСТАННЯ ТЕХНОЛОГІЙ BIG DATA У СИСТЕМІ ОХОРОНИ ЗДОРОВ'Я

У ХХ ст. прагнення людства до забезпечення економічного і технологічного розвитку, як правило, забезпечувалося за рахунок людських можливостей та нещадної експлуатації навколишнього середовища. Наслідки цього у ХХІ ст. є очевидними: світ перебуває у глибокій кризі, а суперечності між соціально-економічними та природними системами досягли загрозливого рівня. Ці обставини змусили прогресивну міжнародну громадськість започаткувати так звану «концепцію сталого розвитку» (sustainable development), як новий підхід до подолання зазначених проблем [1].

Важливою передумовою на шляху втілення концепції сталого розвитку є розвиток системи охорони здоров'я. Нещодавня пандемія виявила ряд недоліків традиційної системи охорони здоров'я та продемонструвала нездатність окремих її одиниць швидко надавати жителям висококваліфіковану медичну допомогу.

Однією з нових технологій, яка допомагає сучасним медичним закладам моніторити стан здоров'я пацієнта, є використання різних смарт-пристроїв: телефонів, годинників, браслетів, окулярів, спеціального взуття та одягу. Розвиток Інтернету речей (Internet of Things) дозволив цим пристроям збирати дані величезного масштабу. Крім того, система охорони здоров'я історично накопичила великі обсяги даних, наприклад, медичні звіти, історії хвороб, рецепти, медичні зображення та різні адміністративні дані. І хоча більшість з них зберігається у паперовій формі, нинішня тенденція спрямована на швидке їх оцифрування.

Особливістю зібраних системою охорони здоров'я даних є значний обсяг, різноманітність типів даних та швидкість їх обробки. Тому аналіз та синтез таких даних є актуальною задачею, розв'язання якої дозволить постачальникам медичних послуг та іншим зацікавленим сторонам в системі надання медичної допомоги виявити необхідні закономірності та тенденції і розробити нові методи діагностики та лікування, які дозволять надати пацієнтам швидку допомогу, що матиме належну якість при порівняно низьких витратах.

Разом з тим специфіка великих даних (Big Data) у галузі охорони здоров'я робить складною (або неможливою) їх обробку за допомогою традиційних інструментів і методів управління даними (рис. 1) або за допомогою традиційного програмного та/або апаратного забезпечення. Тож на сьогодні існують значна кількість інструментів для роботи з великими даними, згрупованих у вигляді табл. 1 [2].

Потенціал аналітики великих даних у сфері охорони здоров'я досить великий. При цьому досягнення кращих результатів можна реалізувати за допомогою різних сценаріїв:

- аналіз характеристик пацієнтів, вартості та результатів лікування з метою визначення клінічно та економічно ефективних методів лікування;
- застосування розширеної аналітики до профілів пацієнтів (наприклад, сегментація та прогнозне моделювання) для визначення осіб, яким буде корисна профілактична допомога або зміна способу життя;
- широкомасштабне профілювання захворювань для виявлення епідемій та підтримки ініціатив щодо їх запобігання;
- збір та публікація даних про медичні процедури, які інформують лікаря щодо ефективності протоколів лікування;
- ідентифікація пацієнтів, які готові прийняти участь в клінічних випробуваннях;
- розробка і впровадження мобільних додатків, які допомагають пацієнтам здійснювати вибір лікаря, лікарям відстежувати дотримання режимів лікування, а системі охорони здоров'я у цілому виявляти тенденції, які призводять до покращення здоров'я населення.

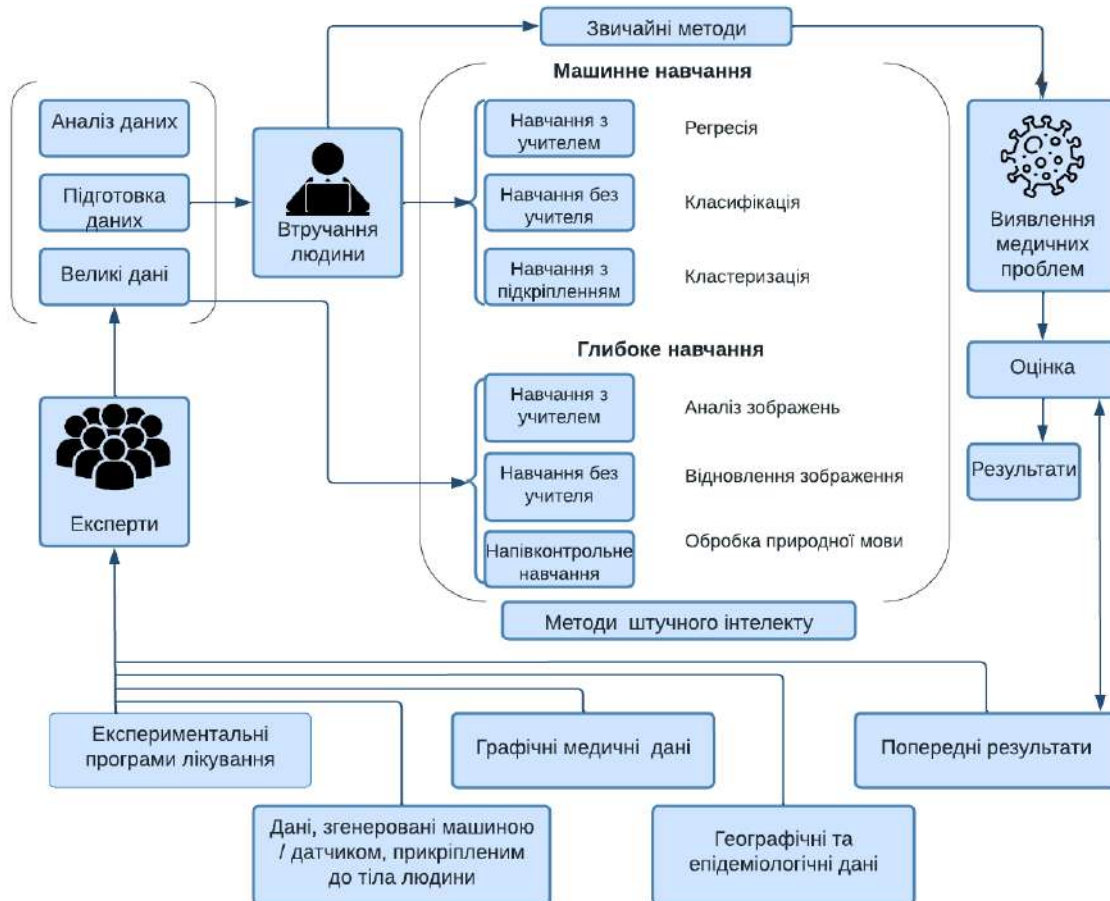


Рис. 1. Методи та підходи обробки великих даних у медичній сфері

Таблиця 1

Інструменти для роботи з великими даними

Інструмент	Опис
Hadoop https://hadoop.apache.org	Інструмент розподіленої обробки даних
Hpc https://hpcsystems.com	Інструмент, який дозволяє використовувати менше коду для комплексної обробки даних
Storm http://storm.apache.org	Інструмент для обробки даних у реальному часі
Qubole https://www.qubole.com	Інструмент, який дозволяє обробляти дані із високим рівнем конфіденційності
Cassandra http://cassandra.apache.org	Розподілена і відмовостійка системи обробки великих даних
Statwing https://www.statwing.com	Інструмент для «очистки» даних та побудови графіків
CouchDB http://couchdb.apache.org	База даних, яка зберігає дані у форматі JSON і надає доступ до даних із використанням HTTP протоколу
Pentaho http://www.pentaho.com	Інструмент для побудови графіків та звітів
Flink https://flink.apache.org	Інструмент обробки потоків даних із здатністю відновлюватися після збоїв
Cloudera https://www.cloudera.com	Платформа для обробки великих даних у різних хмарних провайдерів
OpenRefine http://openrefine.org	Інструмент аналізу великих даних, здатний працювати із великими вибірками і експортувати дані у різні формати
RapidMiner https://my.rapidminer.com	Платформа для підготовки даних для машинного навчання
Hive https://hive.apache.org	Платформа, яка дозволяє виконувати SQL запити до структурованих файлів

І хоча при цьому виникають серйозні юридичні та етичні проблеми, пов'язані з доступом до медичних даних, ми вважаємо, що створення технічних можливостей для доступу до знеособлених агрегованих великих даних та прозорих правил їх використання для різних інститутів громадянського суспільства, вчених та бізнесу, у тому числі з комерційними цілями, стане важливим елементом трансформації суспільства до сталого розвитку.

Список використаних джерел

1. Ярошенко О.І. Проблеми та перспективи формування національної стратегії сталого розвитку / Ярошенко О.І. // Матеріали науково-практичної конференції. Сталый розвиток місцевих громад: роль економічної освіти та науки. – Чернівці : Чернівецький нац. ун-т, 2017. – С. 14-15 с.
2. Kumar, Pardeep, Vishal Jain, and Vasaki Ponnusamy. 2021. The smart cyber ecosystem for sustainable development.
<https://www.vlebooks.com/vleweb/product/openreader?id=none&isbn=9781119761679>.