

дозволить підвищити ефективність лікування та зменшити кількість побічних ефектів.

5. Для успішного впровадження персоналізованої медицини необхідна координація зусиль між лікарями, науковцями, фармацевтами та іншими стейкхолдерами.

Отже, впровадження персоналізованої медицини вимагає інтенсивних наукових досліджень, наявності відповідної інфраструктури та спеціалістів, а також координації зусиль між різними стейкхолдерами. Проте, цей підхід має великий потенціал для поліпшення якості та ефективності лікування, що забезпечить покращення здоров'я.

Список літератури

1. Johnson, J. R., & Zopf, D. A. (2020). *Personalized Medicine: The Path Forward. Clinical Pharmacology & Therapeutics*, 107(1), 20-22.
2. Katsnelson, A. (2018). *Personalized medicine pioneer confronts setbacks. Nature*, 563(7733), S58-S59.
3. Collins, F. S., & Varmus, H. (2015). *A New Initiative on Precision Medicine. New England Journal of Medicine*, 372(9), 793-795.
4. Nebert, D. W. (2019). *Personalized Medicine for Real: Beyond Molecular Medicine. Genes & Development*, 33(3-4), 161-163.
5. Collins, F. S. (2010). *Reengineering Translational Science: The Time Is Right. Science Translational Medicine*, 2(18), 18cm17.

У ДК 004.01

*Гуменюк К.В., студентка 3 курсу
спеціальності 122 «Комп'ютерні
науки»
Хмелівський Ю.С., асистент
кафедри інформаційних технологій*

ПОРІВНЯННЯ ІЄРАРХІЧНОЇ КЛАСТЕРИЗАЦІЇ ТА МЕТОДУ К- СЕРЕДНІХ

Донецький національний університет імені Василя Стуса, м. Вінниця

Кластеризація - це техніка інтелектуального аналізу даних. Цей аналіз досліджує закономірні утворення в наборі даних, відомі як кластери. Кластерний аналіз не потребує групування точок даних у заздалегідь визначені групи, що означає, що це метод навчання без вчителя [1]. Алгоритми кластеризації використовуються для впорядкування даних, категоризації даних, стиснення даних і побудови моделей, виявлення відхилень. Існує багато

методів кластеризації, головними є метод k-середніх та ієрархічний кластерний аналіз [2].

Ієрархічний кластерний аналіз - це метод, який створює ієрархію кластерів, рекурсивно об'єднуючи або розділяючи їх на основі їхньої схожості. Він запускається з кожної точки даних як з окремого кластера, а потім об'єднує або розділяє кластери, доки не буде отримано потрібну кількість кластерів. Цей метод створює дендрограму, яка ілюструє ієрархічні відносини між кластерами .

Метод k -середніх [3] - це популярний алгоритм кластеризації, який має на меті розбити набір даних на k різних кластерів. Він починається з випадкового вибору k центроїдів кластерів і присвоєння кожній точці даних найближчого центроїда. Потім центроїди оновлюються ітеративно шляхом обчислення середнього значення точок даних, віднесених до кожного кластера. Цей процес триває до збіжності, в результаті чого утворюється k кластерів з мінімізованою внутрішньокластерною дисперсією.

Ієрархічна кластеризація дозволяє гнучко досліджувати різні ієрархії кластерів, розрізаючи дендрограму на різних рівнях, тоді як метод k-середніх вимагає заздалегідь вказати бажану кількість кластерів. Ієрархічна кластеризація менше залежить від ініціалізації, оскільки вона рекурсивно об'єднує або розбиває кластери. Також цей метод менш чутливий до шуму в наборі даних.

У ієрархічній кластеризації кластери можуть мати різну форму та розмір, оскільки вона може виявляти складні структури. А метод k-середніх припускає, що кластери мають сферичну форму і схожий розмір. Також цей метод чутливий до початкового вибору центроїдів, оскільки різні ініціалізації можуть призвести до різних результатів кластеризації. Проте, метод k-середніх, як правило, більш ефективний з точки зору обчислень, ніж ієрархічна кластеризація, особливо для великих наборів даних, оскільки він передбачає меншу кількість обчислень.

Вибір між ієрархічним кластерним аналізом та k-середніх залежить від конкретної проблеми, специфіки дослідження, характеристик даних та цілей дослідження. Загалом, обидва методи мають свої сильні та слабкі сторони, тому у багатьох випадках можлива їх комбінація.

Список літератури

1. Штовба С.Д., Козачко О.М. *Machine learning: стартовий курс : електронний навчальний посібник*. Вінниця: ВНТУ, 2020. 81 с.
2. Джеймс Г. Уїттон А Хасті Т. Тібшірані Р. *Введення в статистичне навчання з прикладами на мові R*. Пер. з англ. С. Е. Мастіцко - М.: ДМК Пресс, 2017. - 456 с.
3. *Метод K-середніх (K-means) [Електронний ресурс]* – Режим доступу до ресурсу: https://stud.com.ua/10870/marketing/metod_serednih_means