

УДК 004.62

## ЗАСТОСУВАННЯ КЛАСТЕРНОГО АНАЛІЗУ ДЛЯ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ОРГАНІЗАЦІЇ ПРОЦЕСУ ВАКЦИНАЦІЇ ПРОТИ COVID-19

Ляховецький Д.Р.

доцент кафедри СПЗ Комлева Н.О.

Державний університет “Одеська політехніка”, УКРАЇНА

**АНОТАЦІЯ.** У даній роботі систематизовано та на конкретному прикладі продемонстровано способи використання методу кластеризації даних для організації більш ефективного процесу вакцинації проти Covid-19.

**Вступ.** Станом на зараз світ знаходиться в активній фазі пандемії Covid-19. Незважаючи на активний процес вакцинації, що почався близько 6 місяців тому, її темпи є достатньо низькі і зараз у світі вакциновано хоча б однією дозою певної затвердженої вакцини близько 7.8% населення і лише 3,5% населення вакциновані повністю. Тому **метою** даної роботи є систематизація способів використання методу кластеризації даних задля підвищення ступеню повноти корисної інформації в особи, що приймає рішення з організації процесу вакцинації проти Covid-19. Методи кластеризації даних, які можуть бути імплементовані, включають:

- кластеризацію країн за темпами вакцинації як основу для ефективнішої логістики вакцин між країнами і всередині країн;
- визначення більш ефективного розміщення пунктів вакцинації всередині населених пунктів на основі кластеризації за ступенем щільності розміщення населення [1];
- кластеризацію населення і ефективнішого розподілення людей за групами ризику.

**Основна частина.** В рамках даної роботи було систематизовано способи використання методу кластеризації і реалізовано програмне забезпечення, що на конкретному прикладі показує, як особа, що приймає рішення (ОПР) може підвищити кількість відомої їй інформації і прийняти правильніше рішення стосовно вакцинації.

Програма (реалізована за посиланням як проєкт з відкритим кодом - <https://github.com/DanielLyi/clustering-vaccinations>) реалізує процес кластеризації кількості вакцинованого населення різних країн за часткою людей, що отримали хоча б одну дозу вакцини (вісь X) та повністю вакцинованих (вісь Y). Зовнішній вигляд програми показаний на рисунку 1.

Розроблене програмне забезпечення використовує метод k-середніх [2] для кластеризації даних. Програма дозволяє обирати кількість необхідних кластерів, яка у свою чергу є одним із важливих показників для даного методу і може бути визначена “методом локтя” чи завдяки “коефіцієнту силуєту”.

Це ПЗ не є вичерпним прикладом використання методу кластеризації для вирішення проблеми організації більш ефективного процесу вакцинації. Проте результати використання даної програми можуть стати основою для:

- організації більш ефективної кампанії вакцинації між країнами (постачати більшу кількість вакцин у країни у “гірших” кластерах);
- алгоритмів вирахування необхідності відкриття-закриття кордонів в різних країнах;
- організації ефективної системи Covid-паспортів вакцинації (для визначення де вони потрібні, а де ні; кому вони потрібні, а кому ні).

Зокрема, міжнародна програма COVAX [3] може використовувати подібне програмне забезпечення для визначення країн, куди потрібно більш пріоритетно поставляти вакцини. Можна до певного рівня автоматизувати процес логістики, використовуючи різні алгоритми засновані на цих даних. Більш детальну інформацію можна отримати при використанні даних, щодо віку чи професій вакцинованих. Також COVAX може моделювати хід епідемії, змінюючи кількість поставлених вакцин у “негативні” кластери.

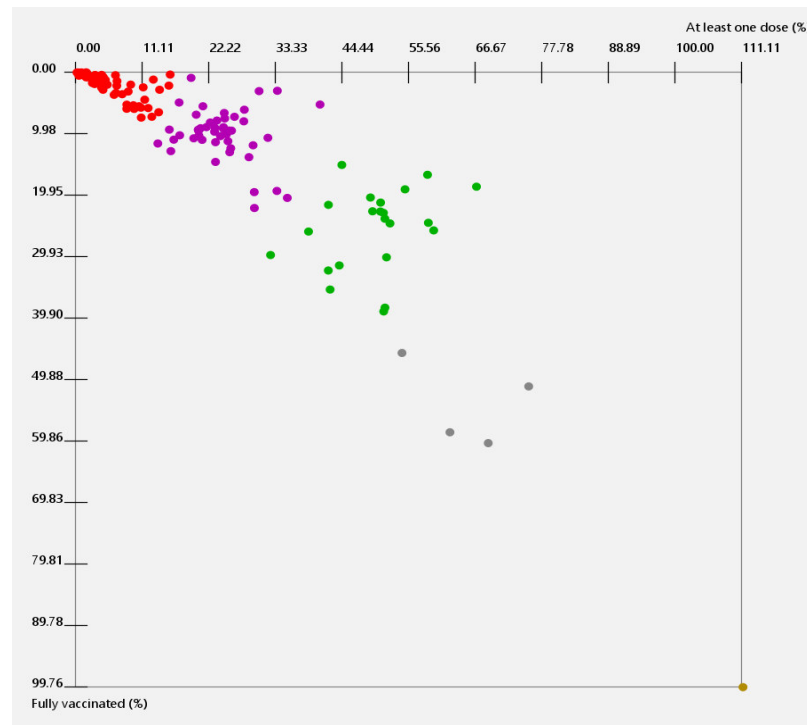


Рис.1 - Графічне представлення результатів застосування програми

Використання методу кластеризації для вирішення проблеми вакцинації проти Covid-19 надає додаткову інформацію  $I_n$  до  $I_{\text{заг}}=I_1+I_2+I_3+\dots+I_n$ . Певне значення  $I_n$  визначається за методом Хартлі та залежить від вхідних даних кластеризації. Тобто таким чином ОПР має додаткові дані для підвищення ефективності програми вакцинації.

Проте, хотілося б водночас наголосити на недоліках даного підходу. По-перше, це важкість визначення необхідної кількості кластерів. Як було зазначено, є різні способи вирішення цієї проблеми, проте вони не надають гарантованого результату. По-друге, можлива невірна інтерпретація даних, що може спричинити неправильні висновки та рішення. Але, незважаючи на ці мінуси, цей підхід може бути поступово застосований і бути використовуваним у процесі вакцинації проти Covid-19.

**Висновки.** Як можна побачити на прикладі розробленого ПЗ, метод кластеризації надає нам додаткові можливості для використання даних для організації більш ефективної програми вакцинації проти Covid-19. Було продемонстровано варіант використання та систематизовано способи використання кластеризації, які можна брати як основу для алгоритмів оптимізації поставки вакцин або для прийняття рішень при поставці вакцин. Тобто, цінність цієї роботи полягає у наданні додаткової систематизованої інформації людям, які можуть використати описані варіанти використання і імплементувати їх, підвищуючи ефективність вакцинації проти Covid-19.

### СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. COVID-19 vaccine: Optimizing cold chain transportation [Online resource]. - Режим доступу: URL: <https://www.geospatialworld.net/blogs/covid-19-vaccine-optimizing-cold-chain-transportation/>
2. The K-Means Clustering Algorithm in Java [Online resource]. - Режим доступу: URL: <https://www.baeldung.com/java-k-means-clustering-algorithm>
3. COVAX. Working for global equitable access to COVID-19 vaccines [Online resource]. - Режим доступу: URL: <https://www.who.int/initiatives/act-accelerator/covax>