

АНОТАЦІЯ

Дипломну роботу виконано на 51 аркушах, вона містить 2 додатки та перелік посилань на використані джерела з 23 найменувань. У роботі наведено 7 рисунків та 11 таблиць.

Метою даної дипломної роботи є створення математичного та програмного забезпечення пошуку інформації серед даних та атрибутів даних у початковому багатовимірному масиву характеристик даних, що розглядаються.

У роботі проведено аналіз існуючих рішень поставленої задачі – методів кластеризації – ієрархічний, k -середніх, c -середніх, expectation-maximization (EM) алгоритм, objective computer clustering (OCC) алгоритм - для даних, регресійного аналізу – лінійний, поліноміальний - для атрибутів даних. Виконано їх порівняння з погляду на складність алгоритмів, їх точність, ефективність та адекватність методів при застосуванні до багатовимірних даних. Для розв'язання задачі в роботі обрано методи кластеризації - ієрархічний, k -середніх, c -середніх – для даних, багатовимірний поліноміальний регресійний аналіз для атрибутів даних.

Для кожного вхідного набору даних сформовано кластеризований вихідний набір даних та регресійна функція. Розроблено автоматизовану систему, що реалізує обрані методи. Виконано тестування розробленої системи.

Ключові слова: багатовимірний масив даних, кластеризація, регресія, інформація, перехресна перевірка, кореляція.