

Фратавчан Тоня Михайлівна, к. ф.-м. н, доцент,
Фратавчан Валерій Григорович, к. ф.-м. н, доцент,
Антонюк Світлана Володимирівна, к. ф.-м. н, доцент
Івасюк Роман Вікторович, аспірант

Чернівецький національний університет імені Юрія Федьковича, Чернівці

<https://orcid.org/0000-0003-1076-0794>

ОСОБЛИВОСТІ КЛАСТЕРИЗАЦІЇ У ВИПАДКУ БАГАТОМОДАЛЬНОСТІ АБО СКЛАДНОЇ ТОПОЛОГІЇ КЛАСТЕРІВ

Задача кластеризації полягає у розбитті загальної множини об'єктів на підмножини таким чином, щоб значення метрики відмінності між екземплярами кожної підмножини були малими, а для екземплярів різних множин ці метрики були великими. У випадку опису об'єктів багатовимірними числовими характеристиками такими метриками можуть бути евклідові міри.

Найпопулярнішими та ергономічними методами кластеризації вважаються метод k -середніх та метод k найближчих сусідів. Але ці методи забезпечують прийнятні результати кластеризації, якщо області локалізації класів у n -вимірному просторі мають форму, наближену до сфероїдальної (рис.1).

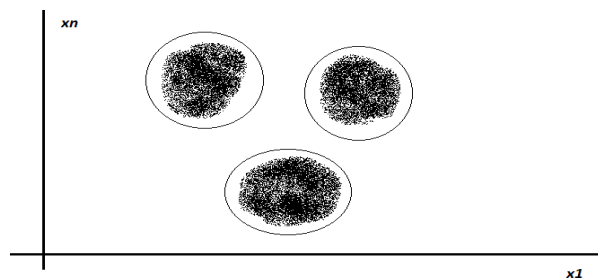


Рис.1. Кластери «сфероїдальної» локалізації.

У випадку більш складної топологічної форми результат застосування цих методів може бути незадовільним (рис.2).



Рис.2. Некоректна кластеризація методом k -середніх.

Для кластеризації множин з багатомодальними класами та класами з більш складними топологічними формами пропонується застосувати апроксимаційні форми Ерміта. Ідея полягає у «прив'язці» кожного класу до деякої багатовимірної параметричної кривої, яка хоча б частково повторювала б конфігурацію самого класу (рис.3).

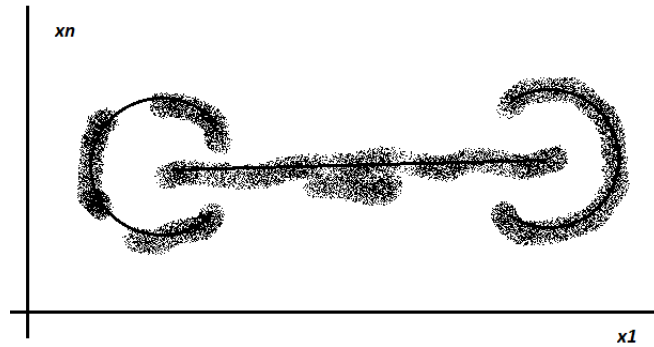


Рис.3. Кластери багатомодальної і складної топологічної структури та їх апроксимація параметричними кривими.

Зручними для програмної реалізації є апроксимаційні форми Ерміта:

$$p(t) = (2t^3 - 3t^2 + 1)p_0 + (t^3 - 2t^2 + t)m_0 + (-2t^3 + 3t^2)p_1 + (t^3 - t^2)m_1 ,$$

де $t \in [0, 1]$ – параметр кривої,

p_0, p_1 – початкова та кінцева точка параметричної кривої,

m_0, m_1 – початковий та кінцевий напрямний вектор.

Обчислення координат початкових та кінцевих точок, а також початкових та кінцевих векторів проводиться як знаходження розв'язку оптимізаційної задачі.

Література

1. Valerii FRATAVCHAN, Tonia FRATAVCHAN. One Pattern Recognition Method for Complex Geometric Clusters Configuration. Proceedings of the 14th International Conference on Development and Application Systems, DAS 2018. (24-26, May 2018, Suceava – Romania), pp.200-203. URL: <http://www.dasconference.ro/dvd2018/data/papers/D51-paper.pdf>.
2. Фратавчан В.Г., Фратавчан Т.М., Лукашів Т.О., Літвінчук Ю.А. Методи та системи штучного інтелекту: навчальний посібник. Чернівці: ЧНУ, 2023, – 115 с.