

## ОГЛЯД МЕТОДУ ГЛИБОКОЇ НЕЧІТКОЇ КЛАСТЕРИЗАЦІЇ ДАНИХ

Таняньський О.С.

Науковий керівник – к.т.н., доц. Шафроненко А.Ю.

Харківський національний університет радіоелектроніки, каф. ІНФ,  
м. Харків, Україна

e-mail: [oleksii.tanianskyi@nure.ua](mailto:oleksii.tanianskyi@nure.ua)

Deep fuzzy clustering is one of the directions in the development of fuzzy clustering methods. This approach combines the advantages of fuzzy logic and deep learning for efficient data separation and analysis.

The main idea is to use a deep neural network to represent data in a new feature space that takes into account the complex relationships between features and helps preserve the structural features of the data. Deep fuzzy clustering can solve the problems associated with matching high-dimensional and fuzzy data, which are often encountered in real-world problems.

Глибинне навчання (Deep Learning) є підгалуззю машинного навчання, яка фокусується на використанні штучних нейронних мереж з численними шарами (глибокими архітектурами) для вирішення завдань розпізнавання патернів та вивчення складних представлень даних. Основною ідеєю глибинного навчання є автоматичне вивчення представлень за допомогою внутрішніх шарів нейронної мережі, що дозволяє системі ефективно адаптуватися до складних завдань та витягати корисні ознаки з великої кількості даних. Глибинні нейронні мережі використовуються для різноманітних задач, таких як розпізнавання зображень, обробка природної мови, рекомендаційні системи та багато інших. Вони виявляються особливо ефективними у випадках, коли дані мають складну структуру або велику розмірність.

Глибинне навчання часто включає в себе тренування моделей на великих обсягах даних та використання глибоких нейронних мереж з численними шарами для автоматичного виділення важливих ознак. Також популярними є концепції, такі як згорткові нейронні мережі (Convolutional Neural Networks, CNN) для обробки зображень та рекурентні нейронні мережі (Recurrent Neural Networks, RNN) для роботи з послідовностями даних. Глибинне навчання вважається однією з ключових технологій, що сприяє розвитку штучного інтелекту та вирішенню різних завдань у різних галузях.

Глибока нечітка кластеризація представляє сучасний підхід до вирішення труднощів, пов'язаних з нечіткою кластеризацією в умовах обробки великого обсягу даних високої розмірності та складних розподілів. Цей метод використовує глибокі нейронні мережі для представлення даних у новому просторі ознак.

Основні особливості глибокої нечіткої кластеризації включають:

– представлення у просторі ознак. Дані перетворюються за допомогою глибоких нейронних мереж у новий простір ознак, де здійснюється подальший аналіз та кластеризація;

– м'яке розділення даних. Метод дозволяє отримати м'яке розділення даних на кластери, що є важливим у випадках, коли об'єкти можуть належати до декількох категорій одночасно;

– уникнення обмежень нечіткої кластеризації. Глибоке навчання допомагає уникнути обмежень традиційних методів нечіткої кластеризації, особливо при роботі з великим обсягом даних та високою розмірністю;

– автоматичне вивчення ознак. Глибокі нейронні мережі в змозі автоматично вивчати важливі ознаки для кращого представлення даних, що сприяє ефективнішій кластеризації;

– використання сучасних досягнень. Метод користується сучасними досягненнями в галузі глибокого навчання та нейронних мереж, враховуючи їхню потужність та ефективність.

Глибока нечітка кластеризація ставить перед собою завдання подолання обмежень традиційних методів та вирішення проблем, пов'язаних з аналізом великих та складних наборів даних.

У контексті навчання репрезентації вводяться три обмеження для поліпшення придатності представлення для кластеризації. По-перше, для забезпечення ефективного представлення даних в новому просторі ознак застосовується архітектура автокодувальника, щоб гарантувати можливість відновлення вихідних даних шляхом декодування закодованого представлення за допомогою іншої нейронної мережі. По-друге, для оптимального розв'язання завдання кластеризації в новому просторі функцій мінімізується внутрішньокластерна компактність і максимізується міжкластерна роздільність. Нарешті, враховуючи, що дані в одному класі повинні бути близькими один до одного, спорідненість між новими уявленнями налаштовується відповідно до дискримінаційної інформації.

Список використаних джерел:

1. C.-H. Chen, V. S. Tseng, T.-P. Hong, et al., "Cluster-based evaluation in fuzzy-genetic data mining," *IEEE transactions on fuzzy systems*, vol. 16, no. 1, pp. 249–262, 2008.

2. H. Liu, M. Shao, S. Li, and Y. Fu, "Infinite ensemble for image clustering," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, pp. 1745–1754.

3. Shafronenko, A. Y., Kasatkina, N. V., Bodyanskiy, Y. V., & Shafronenko, Y. O. (2023). CREDIBILISTIC ROBUST ONLINE FUZZY CLUSTERING IN DATA STREAM MINING TASKS. *Radio Electronics, Computer Science, Control*, (3), 93–97.