

КЛАСТЕРНИЙ АНАЛІЗ ЗІРКОВОГО НЕБА

Кобзев Д.О., Успенський В.Б.

*Національний технічний університет
«Харківський політехнічний інститут», м. Харків*

Кластерний аналіз є одним із методів статистичного аналізу, що дозволяє розподілити об'єкти на групи за певними критеріями. У середині кожної групи повинні виявитися «схожі» об'єкти, а об'єкти різних групи повинні бути якомога більш відмінні.

У зірковій астрономії кластерний аналіз використовується для виявлення зоряних кластерів та асоціацій, що можуть бути важливими для вивчення розвитку галактик та еволюції зірок. Для його проведення необхідно мати дані про розміщення та характеристики зірок. Ці дані можуть бути зібрані за допомогою телескопів та інших приладів, які дозволяють вимірювати яскравість та температуру зірок.

Методика кластерного аналізу полягає в розподілі зірок за певними параметрами, такими як координати на небосхилі, яскравість, температура, спектральний клас тощо. Далі, за допомогою статистичних алгоритмів, зорі групуються у кластери з подібними характеристиками.

Одним із методів, які ми можемо використовувати у аналізі зіркового неба є метод *k*-середніх (*k*-means): цей метод розділяє зірки на попередньо визначену кількість кластерів, які називаються *k*. Спочатку випадковим чином вибираються *k* точок-центроїдів. Далі кожна точка даних призначається до ближчого до неї центроїду, і центроїд перераховується як середнє арифметичне всіх точок, які були призначені до цього кластеру. Цей процес повторюється до тих пір, поки кластери не стабілізуються.

Інший метод, який можемо використовувати у аналізі зіркового неба є метод *s*-means. Він також розділяє дані на кластери на основі подібності між ними, але кожна точка може належати до кількох кластерів з різною вагою. У цьому методі спочатку випадковим чином вибирається кількість кластерів, які називаються *s*. Кожна точка призначається до кожного кластеру з вагою, яка визначається відстанню між точкою та центроїдом кожного кластеру. Далі центроїди перераховуються на основі середньої ваги кожної точки, що належить до кластеру. Цей процес також повторюється до досягнення стабільності кластерів.

Ще один із підходів до кластерного аналізу полягає у використанні графічних методів, таких як дендрограми, що дозволяють візуалізувати залежності між кластерами та зірковими об'єктами.

Описані методи були відпрацьовані на зображеннях зіркового неба, отримані порівняльні характеристики, проаналізовані обмеження. У подальшому апробовані методи можуть бути застосовані до задач кластеризації зображень взагалі. Це зумовить використання методики в задачах розпізнавання образів і фільтрації, медичної діагностики, детектування рухів, дефектоскопії, встановлення прихованих закономірностей і т.п.