

ИСПОЛЬЗОВАНИЕ МЕТОДА К-СРЕДНИХ ДЛЯ ДИАГНОСТИКИ
СТАДИИ ОНКОЛОГИЧЕСКОГО ЗАБОЛЕВАНИЯ

Федулов Кирилл Андреевич, студент группы 345

Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ»

Целью данной работы является ознакомление с практическим применением кластерного анализа.

Кластеризация — это автоматическое разбиение элементов некоторого множества на группы в зависимости от их схожести.

Метод k -средних считается наиболее популярным ввиду того, что алгоритм разбивает заданное множество объектов на указанное число кластеров.

Работа этого метода разбивается на несколько этапов:

- 1) случайно выбрать k начальных «центров масс» кластеров;
- 2) отнести каждый объект к кластеру с ближайшим «центром масс»;
- 3) пересчитать «центры масс» кластеров;
- 4) проверить критерий останковки, и в случае его не выполнения, вернуться к пункту 2;

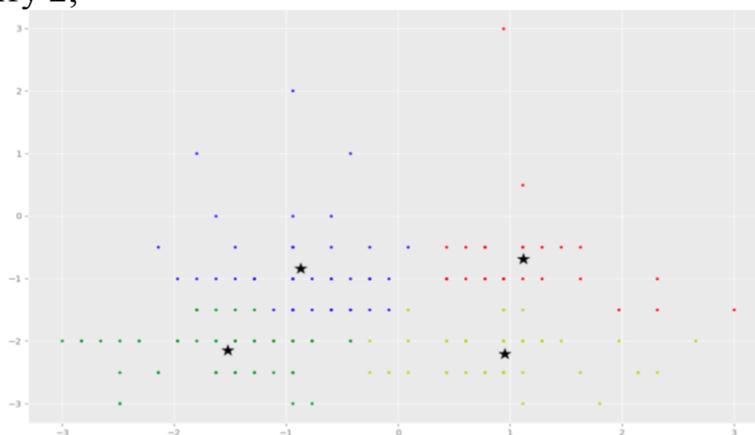


Рис. 1. Возможный результат работы алгоритма

В данной работе был описан алгоритм для решения задачи кластерного анализа, был проиллюстрирован пример работы алгоритма кластеризации (k -средних) с использованием реальных данных об онкобольных пациентах.

Список использованной литературы

1. Воронцов К.В. Алгоритмы кластеризации и многомерного шкалирования [Текст]: учеб. пособие / К. В. Воронцов – Курс лекций. МГУ, 2007г. – 14с.

**Научный руководитель – Меняйлов Е.С., ст. пр. каф. 304.*