**ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ  
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ  
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

Факультет компьютерных наук  
Департамент программной инженерии

|  |  |
| --- | --- |
| СОГЛАСОВАНО  Старший преподаватель департамента программной инженерии факультета компьютерных наук  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Д. В. Пантюхин «\_\_» \_\_\_\_\_\_\_\_\_\_\_ 2017 г. | УТВЕРЖДАЮ  Академический руководитель образовательной программы «Программная инженерия» профессор департамента программной инженерии, канд. техн. наук  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ В. В. Шилов «\_\_» \_\_\_\_\_\_\_\_\_\_\_ 2017 г. |

|  |  |
| --- | --- |
| ***Подп. и дата*** |  |
| ***Инв. № дубл.*** |  |
| ***Взам. инв. №*** |  |
| ***Подп. и дата*** |  |
| ***Инв. № подл*** | RU.17701729.503200-01 34 01-1-ЛУ |

**ПРОГРАММА КЛАСТЕРИЗАЦИИ АЛГОРИТМОМ FASTDBSCAN**

**Руководство оператора**

**ЛИСТ УТВЕРЖДЕНИЯ**

**RU.17701729.503200-01 34 01-1-ЛУ**

Исполнитель  
студент группы БПИ165  
\_\_\_\_\_\_\_\_\_\_ / Соколов Д.В. /  
«\_\_» \_\_\_\_\_\_\_\_\_\_\_ 2017 г.

**Москва 2017**

УТВЕРЖДЕНRU.17701729.503200-01 34 01-1-ЛУ

|  |  |
| --- | --- |
| Подп. и дата |  |
| Инв. № дубл. |  |
| Взам. инв. № |  |
| Подп. и дата |  |
| Инв. № подл | RU.17701729.503200-01 34 01-1 |

**ПРОГРАММА КЛАСТЕРИЗАЦИИ АЛГОРИТМОМ FASTDBSCAN**

**Руководство оператора**

**RU.17701729.503200-01 34 01-1**

**Листов 8**

**Москва 2017**

**Содержание**

[**1. Назначение и область применения** 3](#_Toc451347381)

[**1.1. Назначение программы** 3](#_Toc451347382)

[**1.1.1. Функциональное назначение** 3](#_Toc451347383)

[**1.1.2. Эксплуатационное назначение** 3](#_Toc451347384)

[**1.2. Состав функций программы** 3](#_Toc451347385)

[**2. Условия выполнения программы** 4](#_Toc451347386)

[**2.1. Требования к составу и параметру технических средств** 4](#_Toc451347387)

[**2.2. Требования к информационной и программной совместимости** 4](#_Toc451347388)

[**2.3. Требования к квалификации пользователя** 4](#_Toc451347389)

[**3. Выполнение программы** 5](#_Toc451347390)

[**3.1. Запуск программы** 5](#_Toc451347391)

[**3.2. Ввод входных данных для запуска кластеризации** 5](#_Toc451347392)

[**3.3. Выполнение кластеризации** 6](#_Toc451347393)

[**3.4. Сообщения программы об ошибках** 8](#_Toc451347394)

# **1. Назначение и область применения**

## **1.1. Назначение программы**

Программа «Кластеризация алгоритмом FastDBscan» будет применяться для кластерного анализа входных данных, а также для демонстрации и исследования результата.

### **1.1.1. Функциональное назначение**

Программа будет применяться для кластерного анализа входных данных, а также для демонстрации и исследования результата. У пользователя будет возможность ввести данные точек и один из вариантов кластеризации из доступных, выбрать параметры кластеризации и посмотреть отображаемый результат.

### **1.1.2. Эксплуатационное назначение**

Программа будет использоваться для кластерного анализа входных данных, а также для демонстрации и исследования результата кластеризации, оценки точности, качества и скорости работы различных доступных вариантов кластеризации.

## **1.2. Состав функций программы**

Программа обеспечивает возможность выполнения следующих функций:

1. кластеризация точек из файла с входными данными;
2. настройка параметров для доступных алгоритмов кластеризации;
3. вывод результат кластеризации в форму графика с точками различных цветов, где цвет характеризует принадлежность точки к кластеру;
4. запуск других алгоритмов кластеризации над другими данными и с другими параметрами без перезапуска программы.

# **2. Условия выполнения программы**

## **2.1. Требования к составу и параметру технических средств**

Для нормального функционирования программы требуется компьютер, оснащенный следующими техническими компонентами:

1. процессор не ниже Intel Pentium/Celeron, AMD K6/Athlon/Duron или совместимый с ними с тактовой частотой не ниже 1 ГГц;
2. 512 Мб ОЗУ или более;
3. жесткий диск с объемом свободной памяти не менее 500Мб;
4. VGA-совместимые видеоадаптер и монитор с разрешением не ниже 800х600;
5. клавиатура и мышь.

## **2.2. Требования к информационной и программной совместимости**

Для нормального функционирования программы требуется компьютер, оснащенный следующими программными компонентами:

1. операционная система Microsoft Windows XP (SP2, SP3) / Vista / 7 / 8 / 8.1 / 10;
2. библиотека Microsoft .NET Framework 4.5 и выше.

## **2.3. Требования к квалификации пользователя**

Требуемая квалификация пользователя программы – оператор ЭВМ с базовыми знаниями в области дифференциальных уравнений и нейронных сетей.

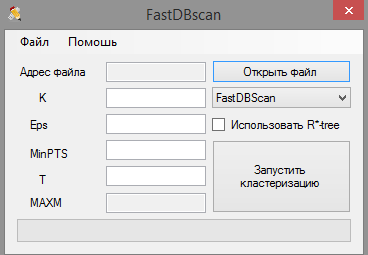
# **3. Выполнение программы**

## **3.1. Запуск программы**

Программа поставляется пользователю на электронном носителе информации в виде исполняемого exe-файла программы и набора dll-файлов с необходимыми для ее работы библиотеками классов.

Программа сразу готова к запуску, ее установка не требуется.

Чтобы запустить программу, необходимо запустить ее exe-файл(Start\_page.exe). В результате на экране появится главное окно программы.



*Рис. 1. Главное окно программы сразу после ее запуска*

## **3.2. Ввод входных данных для запуска кластеризации**

Ввод входных точек для выполнения кластеризации происходит выбором .txt файла в проводнике, который открывается по нажатию на кнопку «Открыть файл». Параметры кластеризации вводятся непосредственно в форму в доступные поля. При наведении на обозначение параметра можно увидеть подсказку по вводу и границам этого параметра.

Адрес файла – адрес .txt файла с точками в заданном формате.

К – количество кластеров для алгоритма кластеризации K-means.

Eps – максимальное расстояние между двумя точками, которые будут считаться плотностно достижымими между собой для алгоритма кластеризации DBscan

minPTS – минимальное количество точек в Eps-окрестности данной точки, при которой она не не будет считаться шумом для алгоритма DBscan.

T – тот процент точек, который берется от полученных кластеров K-means для дальнейшей обработки DBscan’ом при выполнении алгоритма кластеризации FastDBscan

MAXM – то количество точек, которое будет хранится в одном узле R\*-дерева

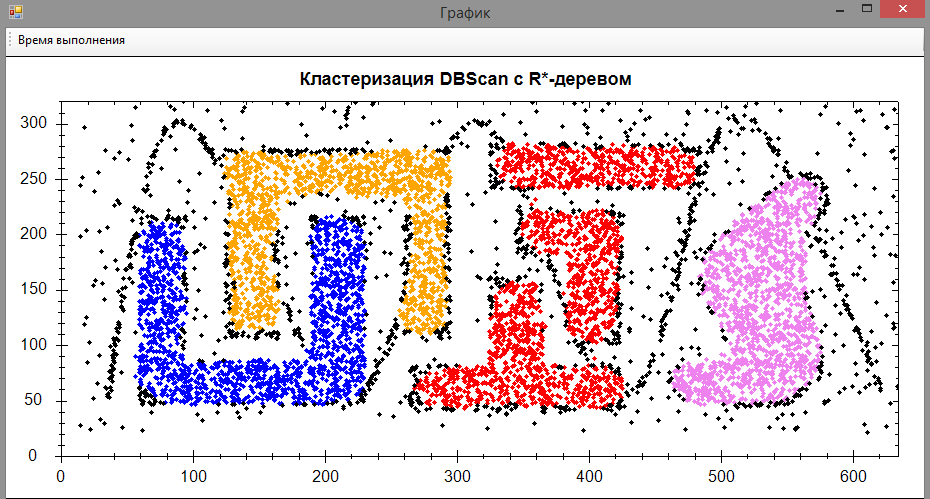
Выпадающий список характеризует вариант выбора доступных алгоритмов кластеризации – DBscan и FastDBscan.

Флажок с надписью «Использовать R\*-tree» обозначает использовать ли в выбранном алгоритме R\*-дерево.

## **3.3. Выполнение кластеризации**

После того, как все необходимые входные данные введены в программу, можно запустить процесс кластеризации. Для этого необходимо нажать кнопку «Запустить кластеризацию» в нижней правой части главного окна программы.

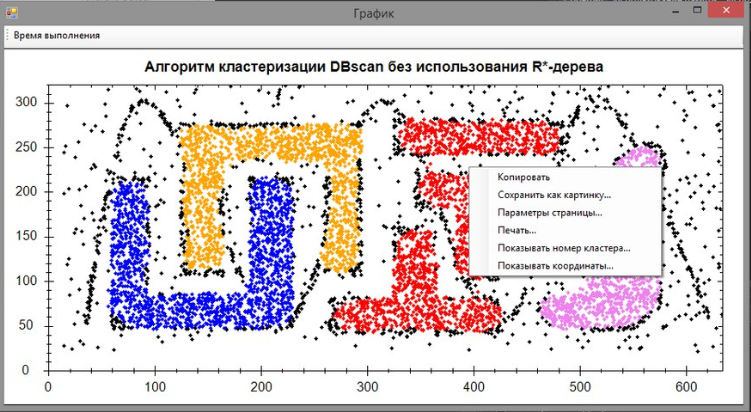
После выполнения процесса решения уравнения будет получено окно с цветовым графиком искомой функции, что и является результатом. Черный цвет на графике и номер кластера “-1” характеризует шум, любой другой цвет и номер кластера – отдельный кластер.



*Рис. 3. Окно с результатами кластеризации (Алгоритм DBscan, Eps=20,minPTS=70,с использованием R\*-дерева)*

По бокам графика находятся координатные оси. В левом верхнем углу находится кнопка «Время выполнения» по нажатию на которую выводятся информация о времени, скорости кластеризации, а также о количестве введенных точек.

При нажатии по графику правой кнопкой мыши появляется варианты работы с графиком.



*Рис. 4. Окно с результатами кластеризации (Алгоритм DBscan, Eps=20,minPTS=70,с использованием R\*-дерева) при нажатой правой кнопкой по графику*

При нажатии на «Копировать» картинка графика сохранится в буфер обмена

При нажатии на «Сохранить как картинку» можно сохранить график ввиде картинки на жестком диске.

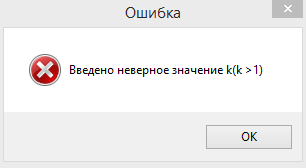
При нажатии на «Параметры страницы» и «Печать» можно изменить настройки печати и распечатать график.

При нажатии на кнопку «Показывать координаты» при наведении на точку на графики будет отображаться ее координаты в формате «Х У» - соответственно абцисса и ордината.

При нажатии на кнопку «Показывать номер кластера» при наведении на точку на графики будет отображаться номер кластера.

## **3.4. Сообщения программы об ошибках**

В случае ввода некорректных входных данных в главное окно программы при попытке нажатия на кнопку «Запустить кластеризацию» появится сообщение с описанием ошибки, а также дополнительная информация по ограничениям параметра. Если ошибка произошла при вводе параметров кластеризации – поле очищается и в поле передается фокус курсора.



*Рис. 5. Пример сообщения программы об ошибке при некорректно введенном K*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Лист регистрации изменений | | | | | | | | | |
| Изм. | Номера листов | | | | Всего листов (страниц) в документе | № документа | Входя- щий № сопро- водит. докум. и дата | Под- пись | Дата |
| изменен- ных | заменен- ных | новых | аннулиро- ванных |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |