Автор:

Mineev S. A. [mineeff20@yandex.ru]

# Добавление новых методов генерации данных по распределениям

Для добавления новых методов генерации данных по распределениям в файле Frameworks/mainwindow.py

1.1 в переменную LIST\_TYPE\_DISTRIBUTION добавить название распределения.

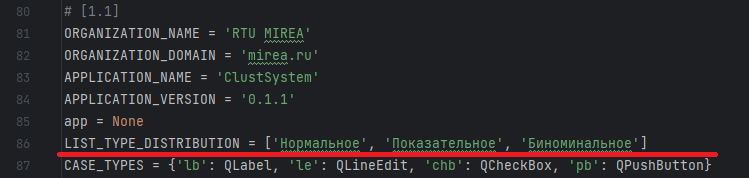


Рис. 1 – Указание названия распределения.

1.2 Определить видимость компонентов для задания параметров в методе

[handler\_pb5\_fr1]. У вас есть два компонента QLabel (labels[0], labels[1]) для отображения названия считываемого параметра и два, соответственно, QLineEdit(line[0], line[1]) Для ввода параметров. Вы можете их видимость отключить(False)/включить(True) с помощью метода setVisible(…). Для изменения названия вы можете использовать методы setText(…), который имеется у labals[0]/[1].

* 1. Определить логику парсинга и занесения в таблицу значение

Для этого у вас имеется метод [handler\_tw1\_fr1], позволяющий представить данные в виде строки параметров. Параметры задаются в виде “key: value” и отделяются между собой “;”.

* 1. В конце нужно реализовать обработчик переданных параметров в методе [handler\_pb6\_fr1]. По умолчанию происходит считывание значений из таблицы и парсинга их. В переменную param1 заносится значение из **line[0]**, а в переменную **param2** значение **line[1]**. Затем они передаются на генерацию данных в соответствующих позициях методов генерации.
  2. Определить методы для генерации в переменной **switch\_case**.

# Добавление новых методов генерации изображений make-\*

2.1 Составить таблицу, чтобы понять какие атрибуты будут использоваться для нового метода напримере таблицы:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| index | type | default\_value [Example] | name\_column | make\_blobs | make\_circles | make\_moons | make\_dna | make\_spheres |
| 1 | int32 | 100 | n\_samples | + | + | + | + | + |
| 2 | int32 | 2 | n\_features | + |  |  |  |  |
| 3 | int32 | None | centers | + |  |  |  |  |
| 4 | float | 1.0 | cluster\_std | + |  |  |  |  |
| 5 | tuple[float, float] | [-0.1,0.1] | center\_box | + |  |  | + |  |
| 6 | int | None | random\_state | + | + | + |  | + |
| 7 | float | 0.8 | factor |  | + |  |  | + |
| 8 | float | None | noise |  | + | + |  | + |
| 9 | float | 1.0 | norm | + | + | + |  |  |
| 10 | float | 0.0 | y | + | + | + |  |  |
| 11 | float | 0.0 | z | + | + | + |  |  |
| 12 | bool | False | shuffle | + | + | + |  | + |
| 13 | bool | False | return\_centers | + |  |  |  |  |

2.2 В случае появления новых параметров добавить значение по умолчанию в переменную DEFAULT\_VALUE и описание в переменную switch\_tootip\_tw2\_fr1.

* 1. Далее вам необходимо определить логику для добавляемого метода в методе cb.currentTextChanged.connect(), который находится в методе инициализации, в кейсе frame1. Согласно таблице вам необходимо у добавляемого метода сделать в таблицы tw1\_fr1 неактивными поля, которым не присвоен знак “+”. Для этого у вас есть переменная ALL, в которой каждому DEFAULT\_VALUE сопоставляется статус отключения с помощью флага-переменной FLAG\_not\_enabled. И для тех полей, у которых установлен в таблицы п. 2.1 знак “+”, установить FLAG\_enabled.
  2. Затем необходимо реализовать логику для переключения между методами на примере “make\_blobs” и на основе данных из таблицы п.2.1 .

Пример:

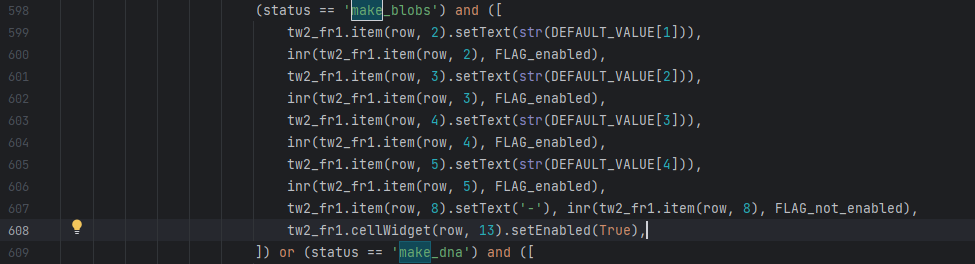


Рис. 2 – Демонстрация установки параметров генерации данных для make\_blobs в таблицу tw2\_fr1 из frame1.

Описание:

Применяем параметры включения/отключения ячеек таблицы при выборе соответствующего метода “make\_blobs” в qcombobox(cb).

* 1. В конце необходимо реализовать логику считывания данных из таблицы и генерации на их основе данных с использованием новых методов make\_\* в методе [handler\_tw2\_fr1].

# Добавление новых методов кластеризации данных.

3.1 Для добавления новых методов кластеризации необходимо определить методы в файле “ ClasteringAlgorithms.py ”, который располагается в директории проекта “ ClusteringMethods”.

Для подключения методов кластеризации используется паттерн проектирования “Стратегия”.

Для добавления нового метода кластеризации необходимо определить класс с названием, которое содержит в имени отсылку на метод, который будет кластеризовывать.

Затем, необходимо реализовать логику двух методов clastering\_points и clastering\_image.

Метод clastering\_points отвечает за кластеризацию данных представленных в виде точек.

Метод clastering\_image отвечает за кластеризацию данных, представленных в виде изображений.

На вход оба принимают данные в виде np.массива координат точек.

Пример:

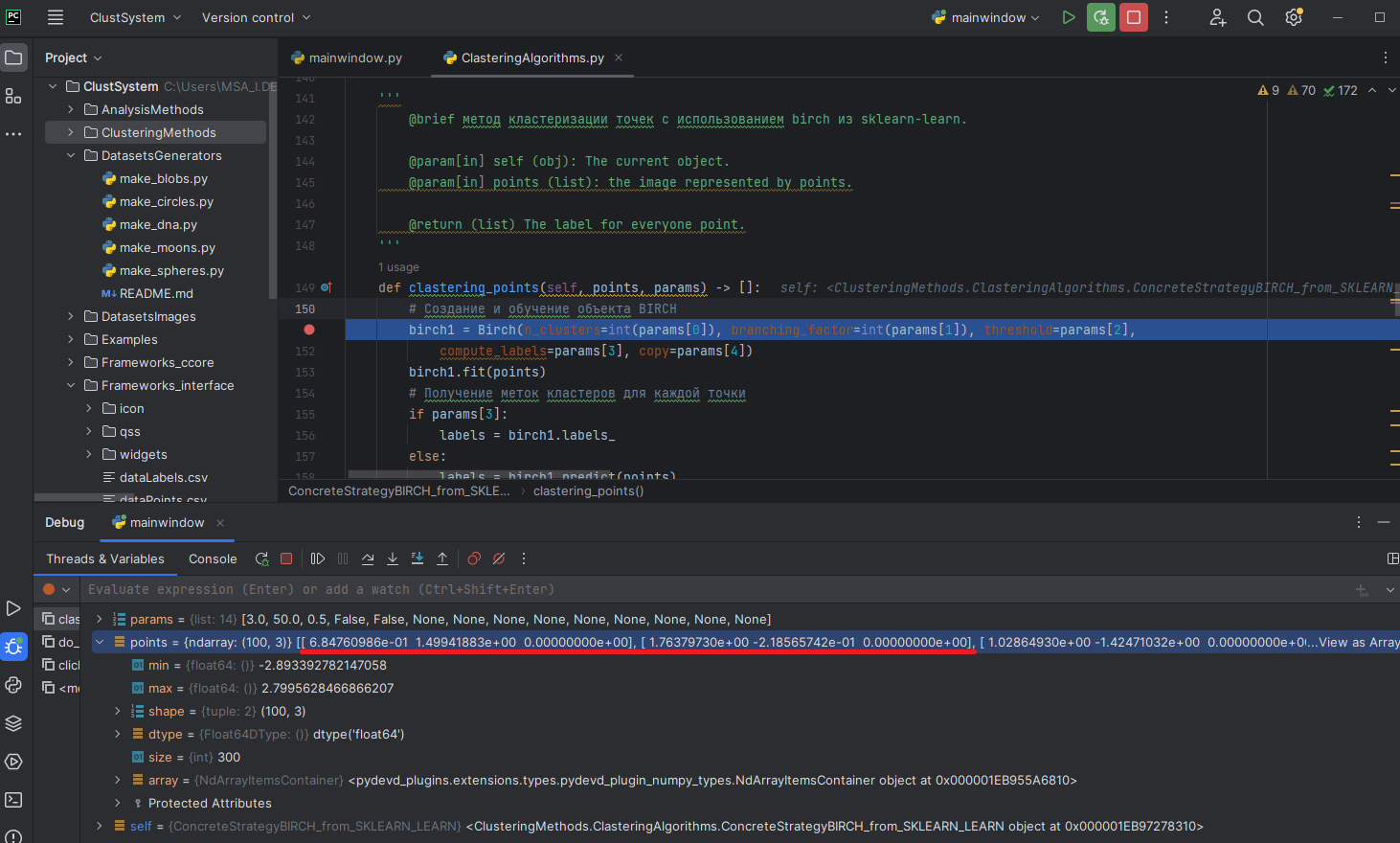


Рис.3 – Демонстрация представления данных в массиве points.

На рис.3 массив points содержит 100 точек, каждая размерностью 3.

На выходе мы получаем метки того, к какому кластеру относится та или иная точка в виде целочисленных неотрицательных чисел, начиная с 0.

3.2 В файле mainwindow необходимо в методе clickStartClustering добавить в раздел кластеризации точек и изображений соответственно контекст метода кластеризации.

* 1. Для того, чтобы иметь возможность задания параметров метода кластеризации нужно реализовать соответствующую логику для [tw3\_fr3]

Необходимо определить в случае отсутствия новых параметров в уже имеющейся таблицы выбора методов кластеризации и задания параметров новые параметры, задав их в опции verticalHeaderLabels компонента tw3\_fr3 = QTableWidget.

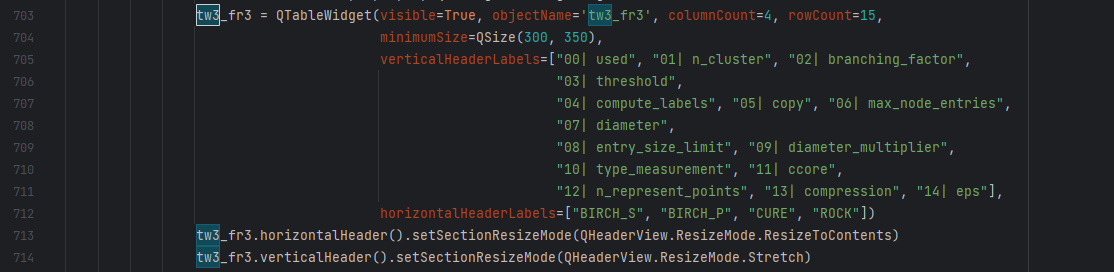


Рис. 4 – Демонстрация параметров таблицы tw3\_fr3.

Затем необходимо реализовать логику добавления компонента в таблицу tw3\_fr3.



Рис. 5 – Демонстрация кода инициализации ячеек таблицы tw3\_fr3.

Описание:

addItemOpen() – Добавляет в i строку, j столбец таблицы tw3\_fr3 компонент QChechBox либо QComboBox в зависимости от того, какой номер строки (i=[0, 4, 5, 11] – методов кластеризации на пересечении в таблицы tw3\_fr3 будет находится компонент в ячейке QChechBox) и (i=[10] – методов кластеризации на пересечении со строкой под индексом 10 будет располагаться компонент QComboBox).

addItemClose() – добавляет в I строку, j столбец таблицы tw3\_fr3 QTableWidgetItem.

В строке 733 кода происходит непосредственное добавление элементов в таблицу, и активация тех ячеек, для которых установлено значение True в переменной switch\_case\_components.