

DOI: 10.33184/dokbsu-2021.5.1

Задача кластеризации в моделировании состава смеси каталитического риформинга бензина

В. А. Алексеева^{1*}, К. Ф. Коледина^{1,2}

¹Уфимский государственный нефтяной технический университет
Россия, Республика Башкортостан, 450064 г. Уфа, улица Космонавтов, 1.

²Институт нефтехимии и катализа УФИЦ РАН
Россия, Республика Башкортостан, 450075 г. Уфа, Проспект Октября, 141.

*Email: valeria.leralex2000@yandex.ru

Статья посвящена исследованию анализа октановых чисел углеводородов процесса каталитического риформинга бензина методами кластеризации в системе Python. Известно, что в каталитическом риформинге присутствует большое количество индивидуальных углеводородов (до 300 разных компонент), что значительно усложняет моделирование. Поэтому их объединяют в группы. Для групп необходимо определять характеристики – октановые числа в первую очередь. Для этого необходимо провести группировку углеводородов по значениям октановых чисел. Для этого будем использовать задачу кластеризации. Задача разбиения заданной выборки объектов на непересекающиеся подмножества, называемые кластерами, так, чтобы каждый кластер состоял из схожих объектов, а объекты разных кластеров существенно отличались. Для решения данной задачи будем использовать задачу кластеризации двумя алгоритмами машинного обучения – это алгоритм иерархической кластеризации и алгоритм K-Means.

Ключевые слова: каталитический риформинг, алгоритмы кластеризации, иерархической кластеризации и алгоритм K-Means.

Каталитический риформинг – современный, широко используемый способ получения высокооктановых бензинов из низкооктанового. Также он необходим для получения ароматических углеводородов (УВ). Важной особенностью процесса является то, что он происходит в среде водорода; его избыток удаляется из системы. Этот водород намного дешевле, чем специально полученный, и он используется в процессах гидрогенизации нефтепереработки.

Задача кластеризации – это обучение без учителя. Требуется разбить конечное множество объектов на кластеры по их взаимной близости. Если для части объектов, обычно очень небольшой, классификации все же известны, то это задача с частичным обучением.

Алгоритмы K-Means направлены на поиск и группирование в классах точек данных, которые имеют высокое сходство между ними. В терминах алгоритма это сходство понимается как противоположность расстояния между точками данных. Чем ближе точки данных, тем больше они похожи и с большей вероятностью будут принадлежать к одному кластеру.

Иерархическая кластеризация – это еще один алгоритм обучения без присмотра, который используется для группировки непомеченных точек данных, имеющих сходные характеристики. Среди алгоритмов иерархической кластеризации различаются два основных типа. Дивизимные или нисходящие алгоритмы разбивают выборку на все более и более мелкие кластеры. Более распространены агломеративные или восходящие алгоритмы, в которых объекты объединяются во все более и более крупные кластеры.

Данные хранятся в файле excel. С колонками: Наименование, Число атомов углерода, Группа, ОЧ ИМХромос (Рис. 1). Всего 318 наименований, 15 чисел атомов углерода и 8 типов таких как: ароматика, изопарафины, нафтен 5-ти членный, нафтен 6-ти членный, оксигенаты, олефины, парафины, нафтены. Данный файл содержит значение октановых чисел (ОЧ) компонентов содержащихся в продукте каталитического риформинга бензинов. Данные могут использоваться для математического моделирования кинетики процесса и определения оптимальных условий его проведения.

1	Наименование	Число атомов углерода	Группа	ОЧ ИМХромос
2	метан	1	Парафины	125,44
3	этан	2	Парафины	125,44
4	пропан	3	Парафины	125,44
5	n-бутан	4	Парафины	113,12
6	n-пентан	5	Парафины	30,85
7	n-гексан	6	Парафины	19,46
8	n-гептан	7	Парафины	15,04
9	n-октан	8	Парафины	32,85
10	n-нонан	9	Парафины	16,1
11	n-декан	10	Парафины	48,49
12	n-ундекан	11	Парафины	48,49
13	n-додекан	12	Парафины	48,49
14	n-тридекан	13	Парафины	48,49
15	n-тетрадекан	14	Парафины	48,49
16	n-пентадекан	15	Парафины	48,49
17	i-бутан	4	Изопарафины	125,44
18	i-пентан	5	Изопарафины	96,25
19	2,2-диметилпропан	5	Изопарафины	215,66
20	2,2-диметилбутан	6	Изопарафины	122,76
21	2,3-диметилбутан	6	Изопарафины	122,76
22	2-метилпентан	6	Изопарафины	108,9
23	3-метилпентан	6	Изопарафины	108,9
24	2,2-диметилпентан	7	Изопарафины	124,14
25	2,4-диметилпентан	7	Изопарафины	124,14
26	2,2,3-триметилбутан	7	Изопарафины	124,14
27	3,3-диметилпентан	7	Изопарафины	43,47
28	2-метилгексан	7	Изопарафины	55,74
29	3-метилгексан	7	Изопарафины	55,74
30	3-этилпентан	7	Изопарафины	90,36
31	2,2,4-триметилпентан	8	Изопарафины	90,36
32	2,2-диметилгексан	8	Изопарафины	34,38
33	2,2,3-триметилпентан	8	Изопарафины	34,38
34	2,5-диметилгексан	8	Изопарафины	34,38

Рис. 1. Значения ОЧ углеводородов.

По данным рис. 1 необходимо сгруппировать УВ по значениям октановых чисел. Провести анализ группировки по классу УВ и числу атомов в структуре молекулы. Сравнить с принятой группировкой. Данные методы и результаты группировки могут в дальнейшем использоваться в моделировании каталитического риформинга бензина.

Для решения данной задачи будем использовать задачу кластеризации двумя алгоритмами – это алгоритм иерархической кластеризации и алгоритм K-Means.

Для определения количества кластеров применен метод локтя. Этот метод позволяет оценить оптимальное количество сегментов. На *рис. 2* приведена дендограмма метода локтя. В данном случае наблюдается, что размер между кластерами не очень большой, есть один большой кластер с 214 элементами. Все остальное соответствует небольшим выбросам или мини кластерам, которые представлены на рисунке 2.

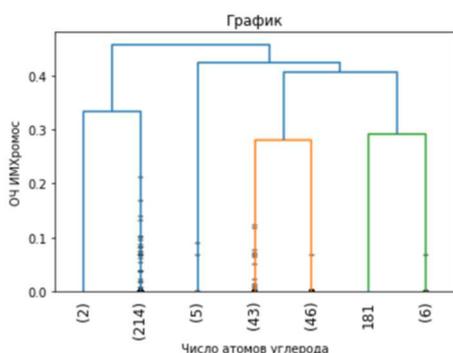


Рис. 2. Дендограмма метода локтя.

На *рис. 3* приведены элементы, попавшие в первый кластер алгоритма K-средних. Здесь элементы с октановым числом равным 48.49 и числом атомов углерода от 10 до 15 (рисунок 3).

Рассмотрим все 5-ти членные нафтены с числом атомов углерода равным 7 (таблица 1).

```
In [26]: df[df['KMeans']==1]
```

Out[26]:

	1Наименование	Число атомов углерода	Группа	ОЧ ИМХромос	Иерархия	KMeans
9	п-декан	10	Парафины	48.49	2	1
10	п-Ундекан	11	Парафины	48.49	2	1
11	п-додекан	12	Парафины	48.49	2	1
12	п-тридекан	13	Парафины	48.49	2	1
13	п-тетрадекан	14	Парафины	48.49	1	1
...
311	И-с11-14	11	Изопарафины	48.49	2	1
312	А-с11-12	11	Ароматика	48.49	2	1
313	1,4-диэтилбензол	10	Ароматика	48.49	2	1
314	А-с10-4	10	Ароматика	48.49	2	1
316	А-с11-15	11	Ароматика	48.49	2	1

Рис. 3. Элементы первого кластера.

Таблица 1. Результат решения задачи кластеризации для 5-ти членных нафтен с числом атомов углерода равным 7

Наименование УВ	Число атомов углерода	Группа	ОЧ ИМХромос	Иерархия	KMeans
этилциклопентан	7	Нафтен 5-ти членный	34.38	2	3
1t,3-диметилциклопентан	7	Нафтен 5-ти членный	90.36	4	2
1с,3-диметилциклопентан	7	Нафтен 5-ти членный	90.36	4	2
1t,2-диметилциклопентан	7	Нафтен 5-ти членный	90.36	4	2

Проанализировав октановые числа углеводородов процесса каталитического риформинга бензина методами кластеризации в системе Python можно отметить, что такие группы как олефины, ароматика, изопарафины, нафтены 5-ти членные и нафтены 6-ти членные необходимо декомпозировать на 2 или даже 3 класса по октановым числам.

Изначальное количество групп было 135 по делению на классы углеводородов и числу атомов углерода в структуре молекулы, но после проведения алгоритма кластеризации двумя способами: иерархической кластеризацией и K-means, групп или кластеров стало 152. Данная группировка углеводородов может применяться при математическом моделировании кинетики процесса каталитического риформинга бензинов.

Работа выполнена по теме «Разработка новых теоретических подходов и программного обеспечения для моделирования сложных химических процессов и поиска соединений с заданными физико-химическими свойствами» (Регистрационный номер: АААА-А19-119022290011-6).

Литература

1. Iranshahi, D., Amiri, H., Karimi, M. Modeling and Simulation of a Novel Membrane Reactor in a Continuous Catalytic Regenerative Naphtha Reformer Accompanied with a Detailed Description of Kinetics // Energy Fuels. 2013. №27. P. 4048.
2. Белоусов Р. Л., Дрожжин Н. А., Костенчук М. И. Построение нечетких лингвистических переменных с использованием методов кластерного анализа данных // Журнал Прикладная информатика. 2015. Т. 10. №1. (55). С. 98–105.
3. Русановский Е. С. Установка каталитического риформинга. -- М.: Химия, 1975. -- 72 с.
4. Суханов В. П. Каталитические процессы в нефтепереработке. М., «Химия», 1973.
5. Гитис, Леонид Хаскелевич. Статистическая классификация и кластерный анализ / Л. Х. Гитис; Моск. гос. гор. ун-т. – М. : Изд-во Моск. гос. гор. ун-та, 2003. – 156, [1] с.
6. Дюран Б. Оделл П. Кластерный анализ. М., «Статистика», 1977. 128 с.

Clustering problem in modeling the composition of a mixture of catalytic reforming of gasoline

V. A. Alekseeva^{1*}, K. F. Koledina^{1,2}

¹*Ufa State Petroleum Technological University*

1 Kosmonavtov Street, 450064 Ufa, Republic of Bashkortostan, Russia.

²*Institute of Petrochemistry and Catalysis of the UFIC RAS*

141 Prospekt Oktyabrya, 450075 Ufa, Republic of Bashkortostan, Russia.

**Email: valeria.leralex2000@yandex.ru*

The article is devoted to the study of the analysis of octane numbers of hydrocarbons of the process of catalytic reforming of gasoline by clustering methods in the Python system. It is known that a large number of individual hydrocarbons (up to 300 different components) are present in catalytic reforming. Which significantly complicates the simulation. Therefore, they are grouped together. For groups, it is necessary to determine the characteristics – octane numbers in the first place. To do this, it is necessary to group hydrocarbons according to the values of octane numbers. To do this, we will use the clustering task. The task of splitting a given sample of objects into disjoint subsets, called clusters, so that each cluster consists of similar objects, and the objects of different clusters differ significantly. To solve this problem, we will use the clustering problem with two machine learning algorithms, this is the hierarchical clustering algorithm and the K-Means algorithm.

Keywords: catalytic reforming, clustering algorithms, hierarchical.