

Билет 17

1. Дайте определение понятия «кластер».

Кластер (англ. cluster — скопление, кисть, рой) — объединение нескольких однородных элементов, которое может рассматриваться как самостоятельная единица, обладающая определёнными свойствами.

2. Охарактеризуйте два любых алгоритма кластеризации.

1. Иерархические и плоские.

- Иерархические алгоритмы (также называемые алгоритмами таксономии) строят не одно разбиение выборки на непересекающиеся кластеры, а систему вложенных разбиений. Т.о. на выходе мы получаем дерево кластеров, корнем которого является вся выборка, а листьями — наиболее мелкие кластера.
- Плоские алгоритмы строят одно разбиение объектов на кластеры.

2. Четкие и нечеткие.

- Четкие (или непересекающиеся) алгоритмы каждому объекту выборки ставят в соответствие номер кластера, т.е. каждый объект принадлежит только одному кластеру.
- Нечеткие (или пересекающиеся) алгоритмы каждому объекту ставят в соответствие набор вещественных значений, показывающих степень отношения объекта к кластерам. Т.е. каждый объект относится к каждому кластеру с некоторой вероятностью.

Практика

```
class DataAnalysis:
    ...
    Определить длину строки.
    ЕСЛИ длина строки превышает 4 символа, ТО вывести строку в нижнем регистре.
    Заменить в строке первый символ на 'O'. Результат вывести на экран.
    ...

    def __init__(self, string=None, firstLetter=None):
        self._string = string
        self._firstLetter = firstLetter
        self._replaceFirstLetter()

    def _replaceFirstLetter(self):
        self._string = self._string[:0] + self._firstLetter + self._string[0 + 1:]

    @property
    def string(self):
        return self._string.lower() if self.len > 4 else self._string
```

```
@property  
def len(self):  
    return len(self._string)
```

```
task = DataAnalysis("У лукоморья 123 дуб зеленый 456", '0')  
print(task.string)
```