**САНКТ-ПЕТЕРБУРГСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ**

**УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ**

**Кафедра Образовательных Технологий**

Дисциплина: Теория искусственного интеллекта

Лабораторная работа №7

Выполнил: Эмрих Н.А.

Группа P3420

Проверила: Жук Ю.А.

Санкт-Петербург

2017

Задание:

В таблице (вариант 2) приложения №2 удалить столбец «группа», выполнить процедуру кластеризации методом *k-means* и построить дендрограмму по оставшимся данным.

Определить ошибку кластеризации, сравнив полученные кластеры с соответствием «группа», т.е. изначальным значением данной графы.

Исходная таблица:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Группа** | **Предшественник, баллы** | **Кол-во удобрений, ц на 1 га** | **Прополка, раз** | **Число дней от залива до сброса воды** | **Число дней от косовицы до обмолота** |
| Низкая урожайность | 2,8 | 1,47 | 1,2 | 115 | 8 |
| Низкая урожайность | 3 | 1,23 | 1,3 | 117 | 7 |
| Низкая урожайность | 2,7 | 1,31 | 1,4 | 114 | 9 |
| Низкая урожайность | 3 | 1,5 | 1,5 | 119 | 10 |
| Низкая урожайность | 3,2 | 1,14 | 1,6 | 120 | 7 |
| Низкая урожайность | 2,8 | 1,22 | 1,6 | 121 | 11 |
| Низкая урожайность | 2,7 | 1,3 | 1,3 | 122 | 8 |
| Низкая урожайность | 3,3 | 1,24 | 1,3 | 118 | 10 |
| Низкая урожайность | 2,8 | 1,16 | 1,9 | 119 | 7 |
| Низкая урожайность | 2,7 | 1,22 | 1,6 | 117 | 9 |
| Низкая урожайность | 2,8 | 1,35 | 1,2 | 119 | 10 |
| Низкая урожайность | 2,9 | 1,4 | 1,4 | 115 | 8 |
| Низкая урожайность | 3,1 | 1,36 | 1,3 | 120 | 11 |
| Низкая урожайность | 2,8 | 1,23 | 1,6 | 114 | 10 |
| Низкая урожайность | 2,7 | 1,3 | 1,4 | 118 | 9 |
| Низкая урожайность | 3 | 1,41 | 1,3 | 117 | 8 |
| Низкая урожайность | 2,9 | 1,28 | 1,4 | 120 | 12 |
| Низкая урожайность | 3,1 | 1,36 | 1,2 | 121 | 9 |
| Низкая урожайность | 2,8 | 1,32 | 1,4 | 118 | 7 |
| Низкая урожайность | 2,9 | 1,4 | 1,5 | 118 | 10 |
| Высокая урожайность | 3,2 | 1,05 | 1,5 | 109 | 9 |
| Высокая урожайность | 2,8 | 1,1 | 1,2 | 108 | 10 |
| Высокая урожайность | 2,9 | 1,2 | 1,6 | 118 | 10 |
| Высокая урожайность | 3 | 1,12 | 1,3 | 110 | 14 |
| Высокая урожайность | 3,3 | 1,08 | 1,4 | 112 | 12 |
| Высокая урожайность | 2,7 | 1,13 | 1,5 | 111 | 15 |
| Высокая урожайность | 3 | 1,18 | 1,7 | 112 | 12 |
| Высокая урожайность | 3,1 | 1,22 | 1,3 | 113 | 14 |
| Высокая урожайность | 3,3 | 1,25 | 1,8 | 112 | 13 |
| Высокая урожайность | 2,9 | 1,1 | 1,7 | 113 | 10 |
| Высокая урожайность | 2,8 | 1,2 | 1,8 | 112 | 15 |
| Высокая урожайность | 3,2 | 1,26 | 1,6 | 113 | 9 |
| Высокая урожайность | 3,4 | 1,28 | 1,8 | 110 | 12 |
| Высокая урожайность | 3,5 | 1,22 | 1,9 | 114 | 13 |
| Высокая урожайность | 3 | 1,19 | 1,7 | 108 | 16 |
| Высокая урожайность | 2,8 | 1,29 | 1,7 | 108 | 12 |
| Высокая урожайность | 2,9 | 1,24 | 1,6 | 112 | 10 |
| Высокая урожайность | 3 | 1,17 | 1,8 | 114 | 9 |
| Высокая урожайность | 3,3 | 1,25 | 1,3 | 115 | 11 |
| Высокая урожайность | 3,4 | 1,27 | 1,7 | 112 | 12 |
| Высокая урожайность | 3,5 | 1,26 | 1,5 | 111 | 14 |
| Высокая урожайность | 3,1 | 1,3 | 1,5 | 119 | 11 |
| Высокая урожайность | 3,2 | 1,24 | 1,6 | 110 | 13 |



Листинг программы:

Вывод в консоль:

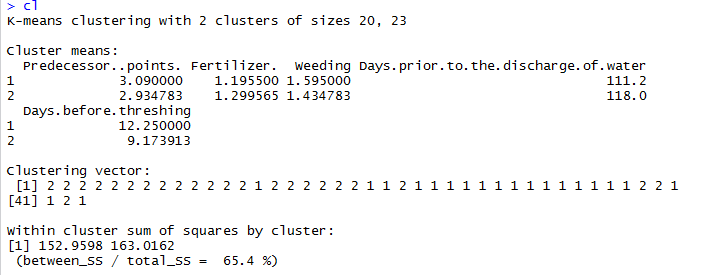
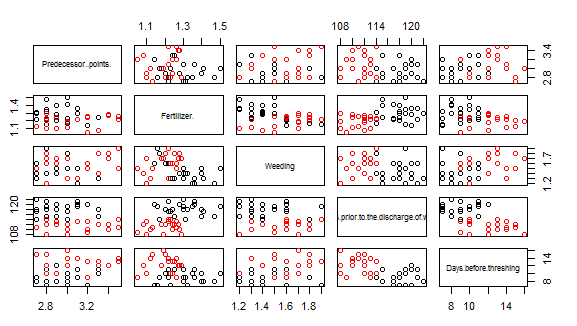
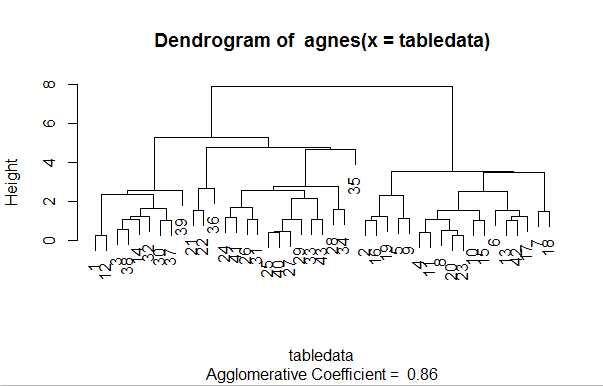


График кластеризации методом *k-means.*



Дендограмма



В нашей данной таблице имеется с 1 по 20 – низкоурожайных данных и с 21 по 43 – высокоурожайных данных. Из дендограммы видно что в левой части собраны элементы таблицы высокоурожайной группы, а в правой низкоурожайной. Посчитаем ошибку кластеризации:  
 Число неверно определенных элементов в левый кластер – 6.

Число неверно определенных элементов в правый кластер – 2.

Общее количество – 43.

Ошибка кластеризации: