**Описание 5 работы - Кластеризация данных**

**Ход работы:**

1. Сформировать датасет (15 столбцов и более)
2. Выполнить кластеризацию
3. Оценить качество кластеризации
4. Выбрать наиболее информативные признаки
5. Выполнить кластеризацию
6. Оценить качество кластеризации
7. Сравнить результаты п.3 и п.6. Сделать выводы
8. Обезличить датасет
9. Выполнить кластеризацию
10. Оценить качество кластеризации
11. Сравнить результаты п.3, п.6. и п.10. Сделать выводы

**Методы кластеризации**

1. Односвязывающий метод
2. Иерархическая
3. Алгоритм максиминного расстояния
4. Алгоритм ISODATA
5. Алгоритмы класса ISODATA
6. Алгоритм CURE (Clustering Using REpresentatives)
7. Алгоритм ФорЭл (FOREL)

**Методы выбора наиболее информативные признаков**

1. Компактность (плотность)
2. Разнесенность образов в пространстве характеристик
3. Метод последовательного сокращения (алгоритм Del)
4. Метод последовательного добавления признаков (алгоритм Аdd)
5. Метод случайного поиска с адаптацией (алгоритм СПА)
6. Метод случайного поиска с адаптацией (алгоритм СПА)

**Методы способа измерения расстояния**

1. Евклидово расстояние
2. Квадрат евклидового расстояния
3. Корреляция Пирсона
4. Расстояние Чебышева (Chebychev)
5. Расстояние Минковского (Minkowski)

**Методы оценки качества кластеризации**

1. Индекс Rand
2. Индекс Жаккара
3. Индекс Фоулкса – Мэллова
4. Индекс Phi
5. Компактность кластеров
6. Отделимость кластеров

|  |  |
| --- | --- |
| Группа 215 | |
| Абрахин Егор Дмитриевич st117271@student.spbu.ru | 3,2,3,4 |
| Антонян Арсений Антонович st117248@student.spbu.ru | 1,5,2,2 |
| Аршинов Артём Викторович st117273@student.spbu.ru | 2,3,1,6 |
| Багров Владимир Алексеевич st117262@student.spbu.ru | 4,6,5,1 |
| Бек Вячеслав Александрович st117260@student.spbu.ru | 5,4,4,3 |
| Беляева Ася Павловна st117245@student.spbu.ru | 6,1,2,5 |
| Голуб Глеб Ярославович st117275@student.spbu.ru | 7,3,3,5 |
| Киселев Михаил Васильевич st117281@student.spbu.ru | 1,6,4,3 |
| Кубянин Никита | 3,2,1,1 |
| Лихошерстова Ксения Андреевна st095261@student.spbu.ru | 6,1,5,2 |
| Овчинников Святослав Александрович st117252@student.spbu.ru | 5,5,5,4 |
| Петрова Екатерина Дмитриевнаst117276@student.spbu.ru | 4,4,4,6 |
| Рылова Александра Игоревна st117013@student.spbu.ru | 7,6,3,6 |
| Сергеев Дмитрий Юрьевич st117243@student.spbu.ru | 2,2,2,2 |
| Сериков Кирилл Геннадьевич st117257@student.spbu.ru | 7,5,1,5 |
| Скорынин Павел Алексеевич st117244@student.spbu.ru | 4,1,1,4 |
| Сурин Игнат Сергеевич st117266@student.spbu.ru | 3,3,2,2 |
| Трофимов Иван Александрович st117263@student.spbu.ru | 1,4,3,1 |
| Черевко Максим Евгеньевич st117272@student.spbu.ru | 2,1,4,6 |

|  |  |
| --- | --- |
| Группа 216 | |
| ФИО | Модификация в части |
| Аникин Роман Сергеевич st116634@student.spbu.ru | 1,6,4,3 |
| Асташенкова Милана Константиновна st117259@student.spbu.ru | 3,2,1,1 |
| Барафанов Артём Александрович st117258@student.spbu.ru | 6,1,5,2 |
| Бурганова Рената Максимовна st117269@student.spbu.ru | 5,5,5,4 |
| Горынцев Роман Николаевич st117251@student.spbu.ru | 4,4,4,6 |
| Добрынин Даниил | 7,6,3,6 |
| Зайнуллин Мансур Альбертович st107049@student.spbu.ru | 1,6,4,3 |
| Кривошеин Алексей Евгеньевич st117264@student.spbu.ru | 2,2,2,2 |
| Крылов Александр Сергеевич st117277@student.spbu.ru | 7,5,1,5 |
| Масгутов Михаил Александрович st117274@student.spbu.ru | 4,1,1,4 |
| Осадчев Тимофей Владимирович st117256@student.spbu.ru | 3,3,2,2 |
| Полугрудов Артём Александрович st117270@student.spbu.ru | 1,4,3,1 |
| Пушкарев Никита Павлович st117247@student.spbu.ru | 2,1,4,6 |
| Пчелкова Анна st116032@student.spbu.ru | 2,4,1,6 |
| Седегова Дарья Павловна st117265@student.spbu.ru | 3,2,3,4 |
| Сергеева Александра Алексеевна st117268@student.spbu.ru | 1,5,2,2 |
| Сон Максим Петрович st117250@student.spbu.ru | 2,3,1,6 |
| Сулимов Андрей Станиславович st117261@student.spbu.ru | 4,6,5,1 |
| Уманец Олег st116029@student.spbu.ru | 5,4,4,3 |
| Шарабарин Михаил st122304@student.spbu.ru | 6,1,2,5 |
| Юдинцев Степан Алексеевич st117249@student.spbu.ru | 7,3,3,5 |