МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ» (НИЯУ МИФИ)

Институт Финансовых Технологий и Экономической Безопасности Кафедра Финансового Мониторинга

Математические методы в задачах финансового мониторинга

Лабораторная работа №2

Работу выполнила: студент группы С18-712: Луканов А.

Проверил:

Москва, 2022

Содержание

[1. Постановка задачи 3](#_Toc102330554)

[2. Выполнение работы 4](#_Toc102330555)

# 1. Постановка задачи

1. Выбрать предмет исследования, а также набор показателей, характеризующих данное явление или процесс.
2. По данным провести классификацию муниципальных образований c помощью метода кластерного анализа нейронных сетей.
3. Сравнить классификации, полученные с помощью агломеративных кластер-процедур, метода К-средних и классификацию с помощью нейронных сетей;
4. Дать экономическую интерпретацию результатов классификации.

|  |  |
| --- | --- |
|  | **ОБЩИЕ КОЭФФИЦИЕНТЫ СМЕРТНОСТИ (число умерших на 1000 человек населения)** |
|  | **СРЕДНЕДУШЕВЫЕ ДЕНЕЖНЫЕ ДОХОДЫ НАСЕЛЕНИЯ (в месяц; рублей)** |
|  | **ОБЩАЯ ПЛОЩАДЬ ЖИЛЫХ ПОМЕЩЕНИЙ, ПРИХОДЯЩАЯСЯ В СРЕДНЕМ НА ОДНОГО ЖИТЕЛЯ (на конец года; квадратных метров)** |
|  | **Уровень занятости для людей в возрасте 15-72 лет(процентов)** |
|  | **КОЭФФИЦИЕНТЫ МИГРАЦИОННОГО ПРИРОСТА на 10 000 человек населения** |
|  | **Зарегистрировано преступлений, связанных с незаконным оборотом наркотиков, на 1000 населения** |
|  | **Зарегистрировано преступлений экономической направленности на 1000 человек** |
|  | **ЧИСЛО ПРЕДПРИЯТИЙ И ОРГАНИЗАЦИЙ (на конец года) на 1000 человек** |
|  | **РАСХОДЫ НА ОХРАНУ ОКРУЖАЮЩЕЙ СРЕДЫ(млн руб) на 1000 чел** |
|  | **ВЫБРОСЫ ЗАГРЯЗНЯЮЩИХ ВЕЩЕСТВ В АТМОСФЕРНЫЙ ВОЗДУХ, ОТХОДЯЩИХ ОТ СТАЦИОНАРНЫХ ИСТОЧНИКОВ тысяч тонн** |

# 2. Выполнение работы

Для начала необходимо загрузить данные в статистику. Они уже прошли стандартизацию, поэтому не нужно проводить операцию повторно.

Изображение выглядит как текст, внутренний, снимок экрана

Автоматически созданное описание

Рисунок 1 – Исходные данные для анализа

Обучим нейронную сеть по алгоритму Кохонена. Получили граф с восьмью нейронами (были заданы при вводе параметров). Можно предположить, что в нашей работе должно быть шесть кластеров (количество классов определяется по количеству цветов). Данный вывод совпадает с количеством классов, используемых в методе К-средних, поэтому в дальнейшем будем использовать 6 нейронов.

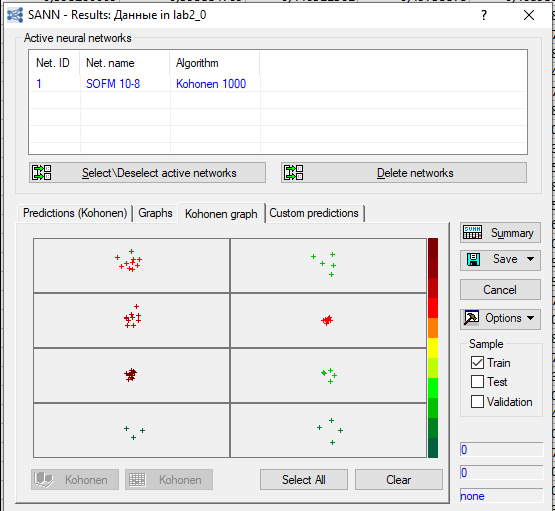


Рисунок 2 – Кластеризация с 8 нейронами

После изменения параметров получили следующий рисунок графа 7 6 нейронами:

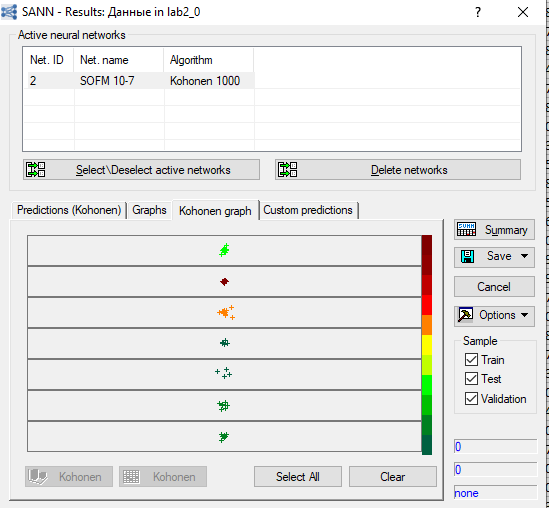


Рисунок 3 – Кластеризация с 7 нейронами

При таком обучении получили следующую матрицу весов:

Изображение выглядит как текст, кроссворд, окно

Автоматически созданное описание

Рисунок 4 – Матрица весов для 7 нейронов

При изменении влияния нейронов на своих соседей при трансформации весов с 3 на 1, получаем следующую матрицу весов. Можно заметить, что веса не сильно различаются и находятся близко друг к другу (сравнивая рисунки 4 и 5). Можно считать веса застабилизировались, поэтому нет необходимости в изменении количества итераций при обучении.

Изображение выглядит как текст, кроссворд

Автоматически созданное описание

Рисунок 5 - Матрица весов для 7 нейронов при изменении влияния нейронов на соседей

Выгрузим данные из статистики в excel и проведем интерпритацию результатов классификации. Сравним функционал качества разбиения.

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 6 - Таблица результатов

Таблица 1 – Распределение объектов по классам

|  |  |  |
| --- | --- | --- |
| Номер кластера | Количество объектов в кластере | Состав класса |
| Кластер 1  S{7} | 7 | г. Москва  г. Санкт-Петербург  Камчатский край  Магаданская область  Ненецкий автономный округ  Сахалинская область  Чукотский автономный округ |
| Кластер 2  S{5} | 5 | Мурманская область  Республика Саха (Якутия)  Тюменская область  Ханты-Мансийский автоном-ный округ – Югра  "Ямало-Ненецкий автономный  округ" |
| Кластер 3  S{9} | 9 | Забайкальский край  "Кабардино-Балкарская  Республика"  Республика Алтай  Республика Бурятия  Республика Дагестан  Республика Ингушетия  Республика Крым  Республика Тыва  Чеченская Республика |
| Кластер 4  S{11} | 11 | Астраханская область  Еврейская автономная область  Карачаево-Черкесская Рес-публика  Курганская область  Омская область  Пермский край  Республика Калмыкия  Республика Коми  "Республика Северная  Осетия – Алания"  Республика Хакасия  Удмуртская Республика |
| Кластер 5  S{5} | 5 | Волгоградская область  Республика Адыгея  Республика Мордовия  Ставропольский край  Чувашская Республика |
| Кластер 6  S{15} | 15 | Алтайский край  Амурская область  Иркутская область  Кемеровская область  Краснодарский край  Красноярский край  Оренбургская область  Приморский край  Республика Башкортостан  Республика Татарстан  Ростовская область  Свердловская область  Томская область  Хабаровский край  Челябинская область |
| Кластер 7  S{33} | 33 | Архангельская область  Белгородская область  Брянская область  Владимирская область  Вологодская область  Воронежская область  г. Севастополь  Ивановская область  Калининградская область  Калужская область  Кировская область  Костромская область  Курская область  Ленинградская область  Липецкая область  Московская область  Нижегородская область  Новгородская область  Новосибирская область  Орловская область  Пензенская область  Псковская область  Республика Карелия  Республика Марий Эл  Рязанская область  Самарская область  Саратовская область  Смоленская область  Тамбовская область  Тверская область  Тульская область  Ульяновская область  Ярославская область |

Таблица 2 – Средние значения показателей

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Кластер 1 | Кластер 2 | Кластер 3 | Кластер 4 | Кластер 5 | Кластер 6 | Кластер7 |
| X1 | -0,65391 | -1,70126 | -1,36532 | -0,12115 | 0,051768 | 0,217372 | 0,702569 |
| X2 | 2,397472 | 1,600437 | -0,71712 | -0,51609 | -0,62419 | -0,07576 | -0,25442 |
| X3 | -0,30635 | -0,90092 | -1,83498 | -0,30803 | 0,054371 | -0,17628 | 0,876501 |
| X4 | 1,643908 | 1,020891 | -1,14694 | -0,81175 | -0,27618 | 0,020255 | 0,112634 |
| X5 | 0,048508 | -0,32409 | -0,05823 | -1,0094 | 0,653046 | -0,04271 | 0,311629 |
| X6 | 0,193778 | -0,22458 | -0,06375 | 1,032475 | -0,74925 | 0,493592 | -0,44469 |
| X7 | 1,188493 | -0,79166 | -0,72395 | 0,838391 | 1,542619 | -0,5059 | -0,21795 |
| X8 | 1,648569 | -0,15849 | -1,02109 | -0,54808 | -0,74642 | 0,107015 | 0,199939 |
| X9 | 0,84644 | 2,962883 | -0,55776 | -0,23505 | -0,48275 | -0,00401 | -0,32304 |
| X10 | -0,40712 | 1,707196 | -0,46694 | -0,25103 | -0,37556 | 0,849092 | -0,29033 |

Рисунок 7 – График средних значений показателей Kohonen

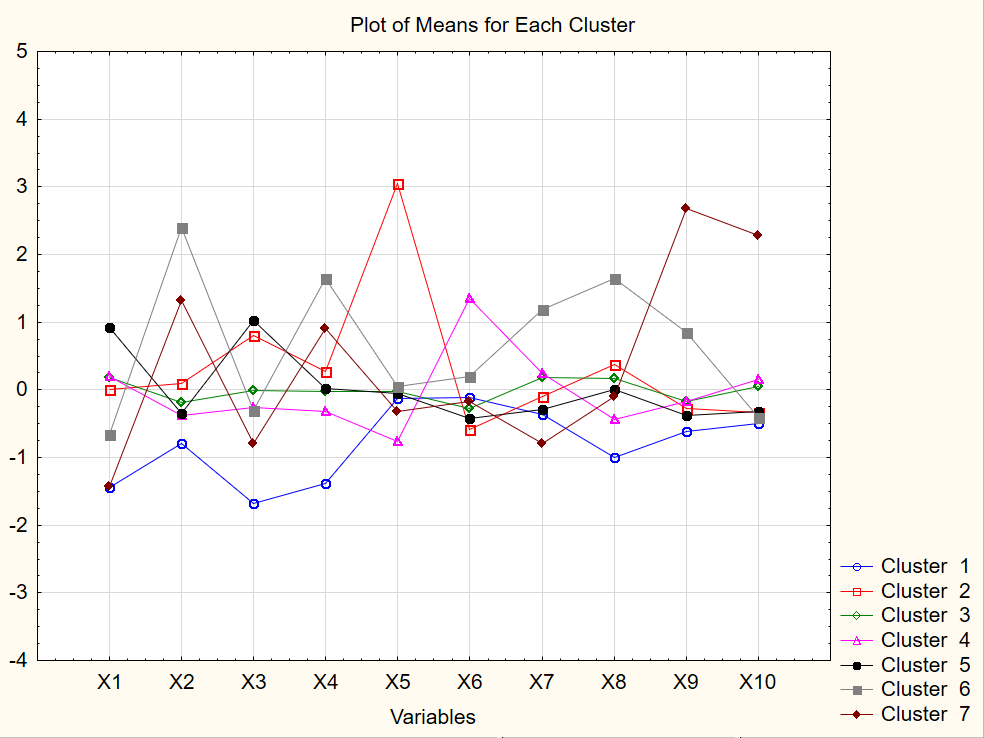


Рисунок 8 – График средних значений показателей k-means

|  |  |
| --- | --- |
| кластер | q |
| 1 | 73,21012 |
| 2 | 47,37738 |
| 3 | 50,74684 |
| 4 | 44,47815 |
| 5 | 20,79356 |
| 6 | 66,3861 |
| 7 | 103,9912 |
| Q | 406,9833 |

Рисунок 9 – Функционал качества разбиения

Выводы:

**Первый класс** состоит из 7 объектов и включает в себя субъекты из центральной России, Севера России и Дальнего Востока. Показатели *среднедушевых доходов, уровня занятости для людей в возрасте 15-72 лет(процентов) и зарегистрированных преступлений экономической направленности на 1000 человек* превалируют по сравнению с остальными классами.

**Второй класс** включает в себя 5 объектов: Мурманская область, Республика Саха (Якутия), Тюменская область, Ханты-Мансийский автономный округ – Югра, Ямало-Ненецкий автономный округ. Характеризуется наибольшими значениями в показателях *расходов на охрану окружающей среды(млн руб) на 1000 чел*. Такие большие показатели можно объяснить тем, что в этот класс входят регионы Сибири.

Кластер также характеризуется самыми низкими значениями *коэффициента смертности, уровня занятости и уровня экономических преступлений.*

**Третий класс** включает в себя 9 объектов и состоит из южных субъектов и регионов Сибири. Характеризуется самыми низкими *среднедушевыми доходами населения, общей площадью жилых помещений, приходящихся на одного человека и числом предприятий и организаций в регионе**на 100 человек.*

**Четвертый класс** состоит из 11 объектов и включает в себя преимущественно субъекты Севера России. Характеризуется наибольшим показателем *зарегистрированных преступлений, связанных с незаконным оборотом наркотиков и наименьшим показателем миграционного прироста на 10000 человек населения*.

**Пятый класс** состоит из 5 субъектов, преимущественно с Юга РФ. Характеризуется самым высоким *коэффициентом миграционного прироста* и *числом* *зарегистрированных преступлений экономической направленности на 1000 человек.*

**Шестой класс** состоит из 15 субъектов, преимущественно из сибирских регионов. Характеризуется средними показателями по всем признакам.

**Седьмой класс** состоит из 33 субъектов, разбросанных по всей России. Характеризуется самыми высокими показателями *коэффициента смертности населения* и *общей жилой площадью, приходящейся в среднем на одного жителя*.