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 №1

Работу выполнила: студент группы С18-702: Луканов А.

Проверил:

Москва, 2022

Содержание

[1. Постановка задачи 3](#_Toc102330554)

[2. Выполнение работы 4](#_Toc102330555)

# 1. Постановка задачи

Ставится задача на основании проведенной кластеризации провести классификацию с помощью нейронных сетей. В качестве исходных данных возьмем результаты кластеризации методом К-средних и разобьем исходные данные на обучающую и тестовую подвыборки.

|  |  |
| --- | --- |
|  | Городское население |
|  | Потребительские расходы на душу населения (тыс.руб) |
|  | Средний размер назначенных пенсий (на 1 января; рублей) |
|  | Численность занятых, приходящихся на одного пенсионера (в среднем за год; человек) |
|  | Число больничных коек (на 1 тыс человек) (на конец года) |
|  | Мощность амбулаторно-поликлинических организаций всего, тыс. посещений в смену (на конец года) |
|  | Прерывание беременности (аборты) на 100 родов |
|  | Заболеваемость на 1000 человек населения |
|  | Численность населения с денежными доходами ниже величины прожиточного минимума в процентах от общей численности населения субъекта РФ |
|  | Общие коэффициенты брачности на 1000 человек населения |

# 2. Выполнение работы

Для начала необходимо определить метод, по которому разобьем на классы наши данные. В нашем случае это будет Метод К-средних.

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 1 – Показатель Q для каждого метода

После определения метода нужно разбить выборку на две части: обучающую и тестовую: для этого разделим исходные данные в отношении 1 к 2.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Название субъекта РФ** | **X1** | **X2** | **X3** | **X4** | **X5** | **X6** | **X7** | **X8** | **X9** | **X10** | **Обучающая выборка** |
| Воронежская область | 0,63 | -0,02 | 1,13 | 0,14 | 0,65 | -0,05 | -0,37 | -0,08 | -0,31 | -0,29 | 5 |
| г. Севастополь | 0,24 | -0,15 | 0,61 | 0,18 | 2,97 | -0,33 | -1,35 | 0,30 | -0,58 | -0,53 | 2 |
| Иркутская область | 0,31 | -0,43 | -0,32 | -0,32 | -0,39 | -0,52 | 0,10 | 0,18 | 0,62 | 0,98 | 3 |
| Калининградская область | -0,18 | -0,24 | 0,64 | 0,76 | 2,26 | -0,69 | -0,57 | 2,41 | -0,58 | -0,48 | 2 |
| Калужская область | 0,81 | -0,06 | 1,18 | 0,76 | -0,34 | 0,13 | 0,35 | 0,25 | -0,40 | -0,47 | 5 |
| Камчатский край | -0,47 | 1,44 | -0,14 | 1,46 | -1,06 | 0,77 | 0,72 | 1,31 | 0,06 | -0,45 | 6 |
| Кировская область | 0,70 | -0,62 | 0,22 | -0,09 | -0,54 | -0,59 | -0,52 | 0,23 | -0,24 | -0,33 | 5 |
| Костромская область | 0,85 | -0,50 | 0,51 | -0,32 | -0,23 | -0,37 | -0,16 | 0,04 | -0,30 | -0,43 | 5 |
| Краснодарский край | -0,01 | 0,24 | 0,20 | 0,10 | 1,05 | -0,26 | 0,04 | -0,14 | -0,50 | 0,48 | 3 |

Рисунок 2 – Обучающая подвыборка

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Название субъекта РФ** | **X1** | **X2** | **X3** | **X4** | **X5** | **X6** | **X7** | **X8** | **X9** | **X10** | **Тестовая выборка** |
| Алтайский край | 0,60 | -0,59 | -0,45 | -0,50 | -0,47 | 0,19 | -1,25 | -0,21 | -0,65 | -0,14 | 3 |
| Амурская область | 0,60 | 0,07 | -0,21 | 0,16 | -0,13 | 2,27 | -0,78 | -0,45 | -0,27 | -0,24 | 4 |
| Архангельская область | 0,24 | 0,24 | 0,51 | -0,63 | -0,62 | -0,97 | 0,61 | -0,43 | 0,10 | -0,06 | 5 |
| Астраханская область | -0,36 | -0,52 | -0,50 | -0,21 | -1,58 | 0,10 | 0,92 | -0,90 | -0,21 | -0,29 | 4 |
| Белгородская область | 0,35 | 0,00 | 1,46 | 0,78 | 0,98 | -1,10 | -0,99 | -0,07 | 0,04 | -0,17 | 5 |
| Брянская область | 0,85 | -0,28 | 1,05 | -0,15 | -0,14 | -0,93 | -0,01 | -0,86 | -0,61 | -0,43 | 5 |
| Владимирская область | 1,17 | -0,49 | 0,92 | 0,57 | 0,24 | -0,87 | -0,01 | -0,01 | -0,52 | -0,41 | 5 |
| Волгоградская область | 0,28 | -0,58 | -0,37 | -0,28 | -0,49 | -1,00 | 0,56 | -0,82 | -0,12 | -0,20 | 5 |
| Вологодская область | 0,63 | -0,28 | 1,21 | -0,36 | -0,47 | -0,91 | -0,42 | 1,15 | -0,12 | 0,44 | 5 |

Рисунок 3 – Тестовая подвыборка

Загрузим обучающую выборку в статистику. Данные уже стандартизированы, поэтому можно пропустить данный шаг.

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 4 – Загруженная выборка для обучения

После загрузки данных можно перейти к обучению персептрона. После обучения моделей можно посмотреть на характеристики предложенных сетей:

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 5 – Характеристики сетей

Здесь мы можем увидеть архитектуру сети, производительность, алгоритм, функцию ошибок и функции активации для скрытых и выходных нейронов. Можно заметить, что 2 сеть обучилась лучше всего.

Можно посмотреть более подробно какие из значений были классифицированы неправильно, а также уровни уверенности для каждой из сетей.

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 6 – Классификация объектов с уровнями уверенности

Проведем классификацию с тестовой подвыборкой с помощью 2 сети и посмотрим на точность классификации.

Изображение выглядит как стол

Автоматически созданное описание

Рисунок 7 – Классификация объектов с помощью обученной сети

Добавляем данные в общий excel-файл и сравниваем результаты с дискриминантным анализом. Сравнивая с дискриминантным анализ, получили результат 0.8. Точность 0.89.