

Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»

Кафедра математической
кибернетики и компьютерных наук

**ПРИМЕНЕНИЕ НЕЙРОННОЙ СЕТИ КОХОНЕНА ДЛЯ
КЛАСТЕРИЗАЦИИ ДАННЫХ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 451 группы
направления 09.03.04 — Программная инженерия
факультета КНиИТ
Тупикова Андрея Юрьевича

Научный руководитель
доцент, к. ф.-м. н.

А. С. Иванова

Заведующий кафедрой
к. ф.-м. н.

С. В. Миронов

Саратов 2016

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1 Сеть Кохонена	5
2 Использование сети Кохонена в программировании. Часть 1.....	6
2.1 Пример работы программы	7
3 Использование сети Кохонена в программировании. Часть 2.....	9
3.1 Пример работы программы	9
ЗАКЛЮЧЕНИЕ	11
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	12

ВВЕДЕНИЕ

В настоящее время все больше возрастает интерес к нейронным сетям. Они применяются повсеместно — бизнес, медицина, физика, техника, геология... Нейронные сети практикуются там, где необходимо решать задачи прогнозирования, классификации и управления. Такое распространение они получили за счет простоты в использовании и богатых возможностей. Нейронные сети — мощный метод моделирования, позволяющий воспроизводить очень сложные зависимости [1], [2], [3].

Нейронная сеть Кохонена используется для решения таких задач, как моделирование, прогнозирование, выявление наборов независимых признаков, но наиболее часто применяется для кластеризации данных. *Кластеризация/классификация* — это разделение исследуемого множества объектов на группы «похожих» объектов, называемых кластерами/классами [4].

Следовательно, тема, рассматриваемая в данной работе является актуальной. Целью работы является реализация сети Кохонена для кластеризации данных. Планируется реализовать два приложения, которые используют способность сети Кохонена классифицировать (кластеризировать) входные данные.

Для реализации поставленных целей используется среда Visual Studio 2010. Разрабатываемые приложения описываются на языке C++.

Первое приложение используется для работы с входными объектами, имеющими по два числовых признака, описывающих данный объект. Это сделано для того, чтобы понять как функционирует нейронная сеть Кохонена и упростить дальнейшее изучение данной сети.

Второе приложение работает с входными объектами, имеющими по три числовых признака. В качестве объектов, в данном случае, будут выступать пиксели какого-либо изображения. Эти три признака у каждого пикселя определяют его цвет в соответствии с цветовой моделью *RGB*.

Данная работа состоит из введения, трех глав, заключения, списка использованных источников и двух приложений. Содержит 34 рисунка, 2 таблицы и 20 использованных источников.

Во введении описаны области применения нейронных сетей и цели работы.

В первой главе даны основные теоретические понятия нейронных сетей, в частности, сети Кохонена. Для описания теоретических данных использовались материалы из [5–17].

Во второй главе рассмотрена работа приложения, использующего объекты, имеющие вектор весов размерностью два. Подробно рассмотрено разработанное приложение, приведены примеры работы.

В третьей главе описана работа приложения, использующего объекты, имеющие вектор весов размерностью три. Также подробно рассмотрено разработанное приложение, в качестве примеров приведены примеры цветного и черно-белого изображений.

В заключении описаны основные выводы.

В приложениях приведены основные функции для работы с разработанными проектами.

1 Сеть Кохонена

Сети Кохонена относятся к самоорганизующимся нейронным сетям. Название самоорганизующаяся сеть означает то, что она обучается без учителя. Нейронная сеть Кохонена позволяет выявлять кластеры (группы) входных элементов, обладающих некоторыми общими свойствами.

Алгоритм функционирования сети:

1. Берем случайным образом один из входных элементов и вычисляем квадрат евклидова расстояния от него до каждого из кластерных элементов сети;
2. Находим минимальное из полученных значений и определяем нейрон-победитель;
3. Для нейрона-победителя, а также для тех нейронов, которые попали в заданный радиус, выполняем корректировку весов;
4. Уменьшаем значение радиуса;
5. Продолжаем обучение, если не выполнено условие остановки обучения.

Остановка обучения происходит в том случае, если величины изменения весов становятся очень маленькими, либо по завершению заданного количества итераций [1], [2], [4], [18–20].

2 Использование сети Кохонена в программировании. Часть 1

Для демонстрации работы нейронной сети Кохонена было написано приложение Windows Forms в среде разработки Visual Studio 2010 на языке C++. Это приложение состоит из нескольких окон, где можно настраивать требуемые для выполнения поставленной задачи параметры: например такие, как число кластеров (нейроны) и объектов (входные элементы). Цель приложения — разбить, используя карту Кохонена, объекты на классы. Структура приложения показана на рисунке 1.

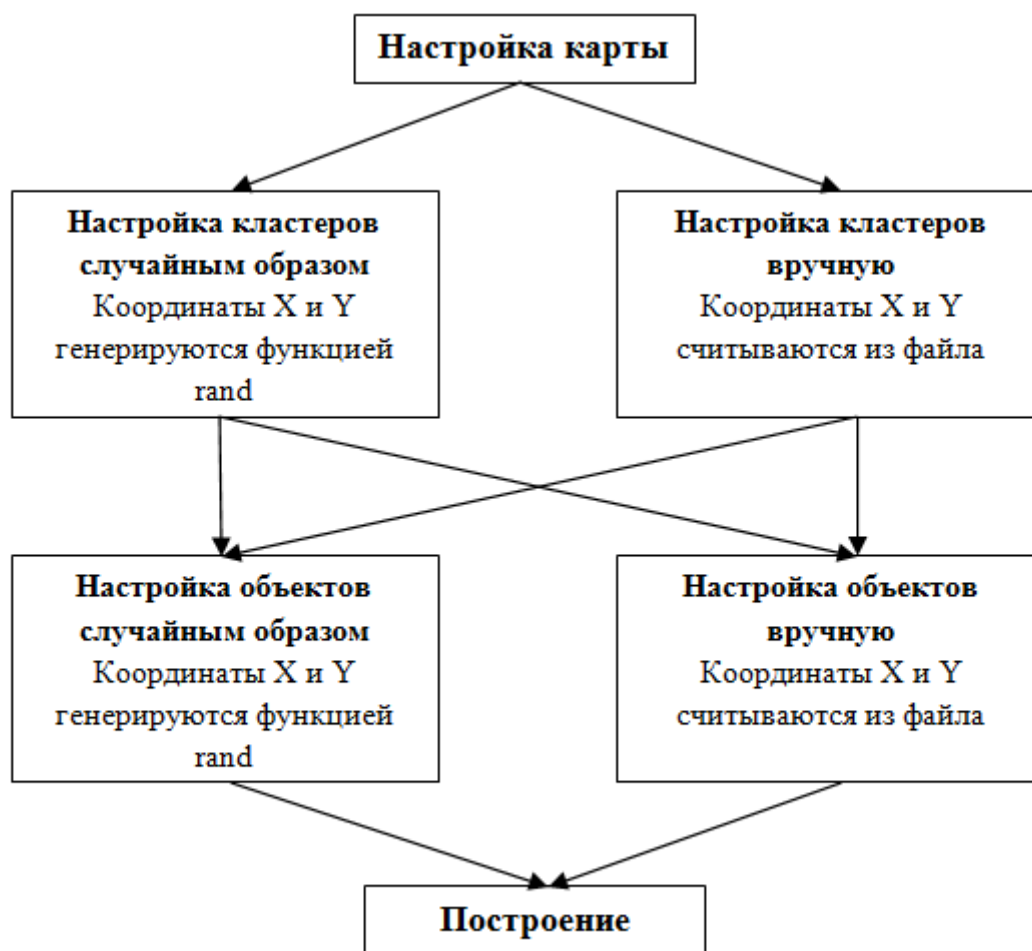


Рисунок 1 – Структура приложения

По данной структуре видно, что приложение также позволяет задавать веса для нейронов и объектов как вручную, так и случайным образом. Каждый объект и каждый кластер имеет свой вектор (в данном случае массив) весов. Размерность такого массива, при задании кластеров и объектов случайным образом ограничена, а при задании кластеров и объектов вручную не ограничена, однако использует приложение всего два веса из списка на выбор

(координаты x и y). Помимо этого, в приложении присутствует ряд исключений: например исключение, которое не позволяет начать обучение сети, если не задана норма, радиус или другой какой-либо важный параметр.

2.1 Пример работы программы

Пусть в некотором городе N некто T решил провести турнир по боксу. Он собрал желающих поучаствовать в этом турнире и их оказалось 15 человек. Перед T встал вопрос о том, как же этих людей грамотно и быстро разделить на группы для проведения боев, чтобы на ринге боролись бойцы примерно равные по силе друг другу. Разработанное приложение позволяет найти ответ на поставленный вопрос.

Нас здесь будут интересовать такие сведения о человеке: его вес (кг) и количество проведенных боев на подобных соревнованиях или турнирах. Сведения о каждом из участников представлены в таблице 1.

Таблица 1 – Сведения о участниках турнира по боксу

Номер участника	Вес (кг)	Количество боев
1	75	23
2	99	56
3	67	12
4	65	50
5	67	23
6	55	12
7	90	0
8	73	2
9	85	13
10	76	8
11	65	9
12	78	46
13	75	1
14	64	20
15	71	34

Для приложения в качестве входных данных будут задаваться участники турнира и сведения о них — вектор весов (вес участника, количество боев). Радиус обучения зададим равным 5. Количество кластеров (нейронов) на самой карте будет равно тоже 5 и вектор весов для них будем задавать случайным образом. Число объектов равно 20 и вектор весов для них будут задаваться вручную (считываться из файла). Число параметров объекта равно 2, норма обучения — 0.1, количество итераций — 10.

На рисунке 2 показан результат работы программы после одной итерации и конечный результат — после 10 итераций.

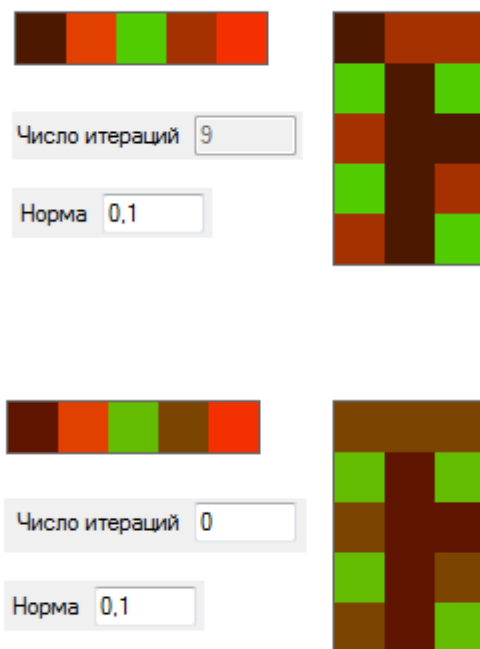


Рисунок 2 – Результат программы для примера с турниром

Как видно, групп получилось три, хотя нейронов было задано пять — это значит, что какие-то нейроны не участвовали в образовании классов. Тогда число нейронов нужно либо сократить при дальнейших опытах с кластеризацией этих бойцов, либо увеличить количество итераций. Однако, в данном случае нас устраивает такой ответ — количество участников в каждой группе больше одного, а значит мы можем разбить всех людей из этих групп по парам для проведения турнира. В первую группу попали участники под номерами: 1, 3, 5, 6, 11, 14. Во вторую: 2, 4, 12, 15. В третью: 7, 8, 13, 9, 10.

Подробное описание работы приложения и примеров работы приведено в бакалаврской работе.

3 Использование сети Кохонена в программировании. Часть 2

Благодаря первому приложению, можно наблюдать в реальном времени как происходит обучение нейронной сети Кохонена. Однако в нем мы использовали вектор весов всего лишь размерностью два. Поэтому, было написано приложение Windows Forms в среде разработки Visual Studio 2010 на языке C++, которое, в отличие от первого, использует вектор весов размерностью три у объектов и кластеров. Особенностью данного приложения является не только размерность вектора, но и задача, которую он выполняет. Это приложение может разбить, используя карту Кохонена, пиксели изображения на группы, что в дальнейшем позволит нарисовать это изображение в контрасте

3.1 Пример работы программы

В качестве примера, было взято цветное изображение размером 133×100 пикселей, показанное на рисунке 3 слева.



Рисунок 3 – Изображение в контрасте цвета для примера с цветным изображением

Его мы задали как входные данные в приложении и использовали такие значения параметров: количество нейронов равно 7, норма обучения — 0.1, радиус — 25 и число итераций — 5.

Нейронная сеть Кохонена разбила входную картинку на четыре группы пикселей. Все нейроны и пиксели, которые относятся к тому или иному нейрону, изображены на рисунке 4.

Здесь мы видим, что карта Кохонена имеет пять разных цветов. Нейроны 2 и 7 имеют одинаковый цвет, соответственно и пиксели, которые относятся к 2 и 7 нейрону мы объединяем в одну картинку, что и видно на рисунке 4. Для нейронов 3, 4 и 5 нет нейронов с таким же цветом, здесь мы просто выводим

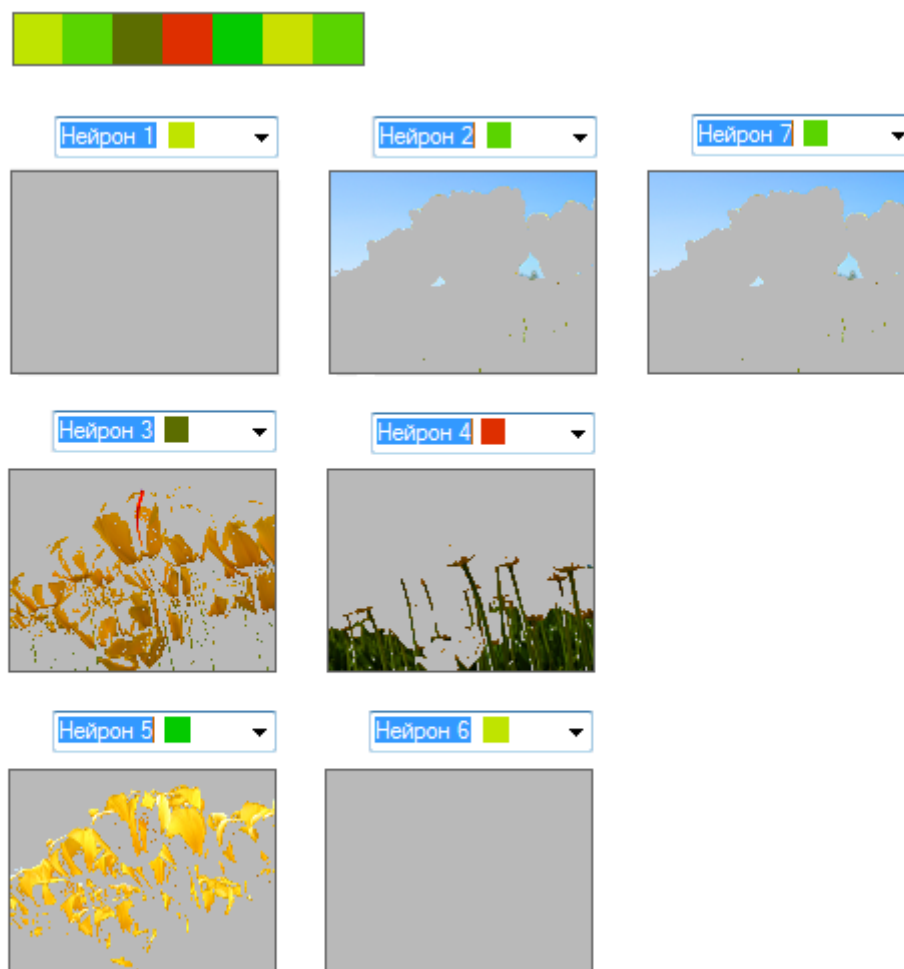


Рисунок 4 – Пиксели, которые относятся к тому или иному нейрону для примера с цветным изображением

те пиксели, которые к ним относятся. А вот нейроны 1 и 6 тоже имеют одинаковый цвет, но ни у 1 ни у 6 нейрона нету пикселей. Все это — нейроны с одинаковыми пикселями и нейроны без пикселей, говорит о том, что либо на карте было задано слишком много нейронов и их количество следует уменьшить, либо нужно увеличить количество итераций. Однако, например, у нейрона 3 есть красный цвет на картинке, что по идее мог бы и быть отдельной группой — это уже нам говорит о том, что нужно либо увеличить количество итераций, либо увеличить количество нейронов на карте Кохонена.

Приложение, опираясь на то, как разбиты пиксели по группам, так же вывело в качестве результата входную картинку, но только в контрасте цвета. Это показано на рисунке 3 справа.

Подробное описание работы приложения и примеров работы приведено в бакалаврской работе.

ЗАКЛЮЧЕНИЕ

В данной работе была рассмотрена одна из моделей нейронной сети — сеть Кохонена. На основе алгоритма функционирования данной сети было написано два приложения Windows Forms. Первое из них работает с входными объектами, имеющими по два числовых признака, а также, позволяет наблюдать в реальном времени процесс обучения сети и процесс формирования групп из исходных объектов. Второе — использует входные объекты, имеющие по три числовых признака, и работает с пикселями любого изображения. Приведены примеры работы разработанных приложений. Показано, что размерность вектора весов и область применения данной сети ограничивается лишь фантазией разработчика.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 *Кохонен, Т.* Самоорганизующие карты, 2-е издание / Т. Кохонен. — Москва: БИНОМ. Лаборатория знаний., 2014.
- 2 Практическое применение нейронных сетей. [Электронный ресурс]. — URL: <http://finteks.ru/prakticheskoe-primenenie-neyronnyih-setey/> (Дата обращения 11.02.2016). Загл. с экр. Яз. рус.
- 3 *Осовский, С.* Нейронные сети для обработки информации / С. Осовский. — Москва: Финансы и статистика, 2002.
- 4 Сети и карты Кохонена. [Электронный ресурс]. — URL: http://gorbachenko.self-organization.ru/articles/Self-organizing_map.pdf (Дата обращения 03.02.2016). Загл. с экр. Яз. рус.
- 5 Нейронные сети. Модели сетей. [Электронный ресурс]. — URL: <http://elanina.narod.ru/lanina/ind/neiro/1.htm> (Дата обращения 17.01.2016). Загл. с экр. Яз. рус.
- 6 *Бодянский, Е. В.* Искусственные нейронные сети / Е. В. Бодянский, О. Г. Руденко. — Харьков: ТЕЛЕТЕХ, 2004.
- 7 *Хайкин, С.* Нейронные сети. Полный курс / С. Хайкин. — Москва: Издательский дом «Вильямс», 2006.
- 8 Классификация известных нейросетей по основным категориям применения. [Электронный ресурс]. — URL: http://www.victoria.lviv.ua/html/oio/html/theme7_1_rus.htm (Дата обращения 21.03.2016). Загл. с экр. Яз. рус.
- 9 Зачем нужны нейронные сети? [Электронный ресурс]. — URL: <http://www.uran.donetsk.ua/~masters/2013/fknt/perevozchikov/library/article6.htm> (Дата обращения 05.02.2016). Загл. с экр. Яз. рус.
- 10 Модель бионической нейронной сети и ее применения. [Электронный ресурс]. — URL: http://www.keldysh.ru/papers/2008/prep89/prep2008_89.html (Дата обращения 1.02.2016). Загл. с экр. Яз. рус.

- 11 Классификация нейронных сетей. [Электронный ресурс].— URL: <http://www.aiportal.ru/articles/neural-networks/classification.html> (Дата обращения 2.02.2016). Загл. с экр. Яз. рус.
- 12 Обзор теории интеллектуального анализа данных на базе нейронных сетей. [Электронный ресурс].— URL: http://www.math.spbu.ru/user/gran/soi11_3/p3-17.pdf (Дата обращения 4.02.2016). Загл. с экр. Яз. рус.
- 13 Биологический нейрон и его кибернетическая модель. [Электронный ресурс].— URL: <http://pandia.ru/text/78/046/86405.php> (Дата обращения 1.02.2016). Загл. с экр. Яз. рус.
- 14 Искусственный нейрон. [Электронный ресурс].— URL: <http://www.trinitas.ru/rus/doc/0016/001d/2430-nik.pdf> (Дата обращения 3.02.2016). Загл. с экр. Яз. рус.
- 15 Физиология человека. [Электронный ресурс].— URL: <http://www.bibliotekar.ru/447/12.htm> (Дата обращения 1.02.2016). Загл. с экр. Яз. рус.
- 16 Биологический прототип. [Электронный ресурс].— URL: http://lms.kgeu.ru/pluginfile.php?file=/52117/mod_resource/content/2/%D0%A0%D0%90%D0%97%D0%94%D0%95%D0%9B1%20%D0%9E%D0%A1%D0%9D%D0%9E%D0%92%D0%AB%20%D0%98%D0%9D%D0%A1.%20%D0%9F%D0%95%D0%A0%D0%A6%D0%95%D0%9F%D0%A2%D0%A0%D0%9E%D0%9D.pdf (Дата обращения 7.02.2016). Загл. с экр. Яз. рус.
- 17 Биологические основы функционирования нейрона. [Электронный ресурс].— URL: http://stu.scask.ru/book_ns.php?id=4 (Дата обращения 5.02.2016). Загл. с экр. Яз. рус.
- 18 Что такое нейронные сети? [Электронный ресурс].— URL: <http://www.help-on-forex.com/strateg148.html> (Дата обращения 07.02.2016). Загл. с экр. Яз. рус.
- 19 Нейронные сети. Основные модели. [Электронный ресурс].— URL: <http://nncourse.chat.ru/course.pdf> (Дата обращения 03.02.2016). Загл. с экр. Яз. рус.

20 Нейронные сети Кохонена. [Электронный ресурс].— URL: <http://neuronus.com/nn/38-theory/961-nejronnye-seti-kokhonena.html>
(Дата обращения 15.03.2016). Загл. с экр. Яз. рус.