МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение

высшего образования

«Вятский государственный университет»

Факультет автоматики и вычислительной техники

Кафедра ЭВМ

Отчёт

Лабораторная работа № 2 по дисциплине

«Системы обработки знаний»

Выполнил студент группы ИВТ-41\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ / Птахова А.М/

Проверил доцента кафедры ЭВМ\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_/ Ростовцев В. С./

Киров 2024

1. Цель

Ознакомиться с основными командами создания, обучения и применения нейронных сетей кластеризации и классификации данных в Neural Network Toolbox с помощью сетей Кохонена и LVQ.

2. Задание

1) Выполнить кластеризации с применением нейронных сетей Кохонена в соответствие с заданием (Приложение А). Создать и выполнить моделирование нейронной Кохонена, согласно согласованному с преподавателем, приведенному в приложении А. Результаты кластеризации вывести в виде графика с разными цветами кластеров, а также точки тестовых векторов.

2) Для создания сети LVQ обучающую последовательность использовать сеть Кохонена.

3. Ход выполнения

3.1 Самоорганизующаяся карта Кохонена

Обучение

Для кластеризации данных был найдена датасет, состоящий из возраста и веса людей. Датасет приведен в приложении А.

Для отображения элементов выборки использовался следующий код:

plot(indata(1,:),indata(2,:),'.g','markersize',20)

hold on

Результат выполнения представлен на рисунке 1.

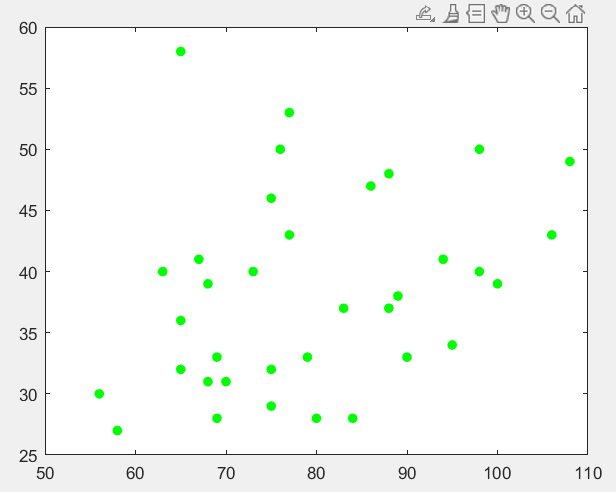


Рисунок 1 – Отображение элементов обучающей выборки

Из-за того, что в обучающей выборке есть похожие данные, то точки накладываются друг на друга.

Для создания самоорганизующейся карты Кохонена выл выполнен следующий код в программе MatLab:

net = selforgmap([6,6]);

Была создана гексагональная сетка размера 6х6, что соответствует 36 нейронам или, условно, кластерам.

Для обучения сети использовалась следующая команда:

[net, tr] = train(net, indata);

Процесс обучения изображен на рисунке 2.

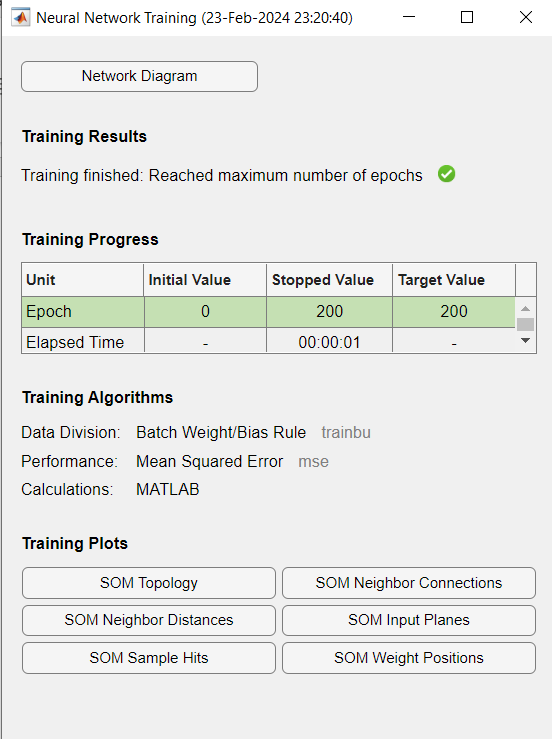


Рисунок 2 – Процесс обучения сети

В результате были получены новые значения весов нейронов. Узлы карты Кохонена были отображена на графике с помощью следующего кода

W = net.iw{1,1}

for i = 1:length(W)

text(W(i,1),W(i,2),int2str(i),'HorizontalAlignment','center', 'VerticalAlignment', 'bottom')

end

hold on

plotsom(net.iw{1,1},net.layers{1}.distances)

Результат выполнения кода представлен на рисунке 2.

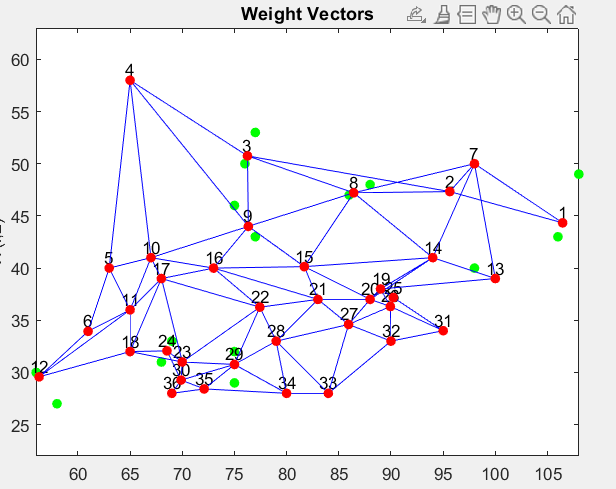


Рисунок 3 – Отображение узлов карты Кохонена после обучения

Как видно из полученного графика, были выделены кластеры, размером с заданной картой. Особо выделаются крайние кластеры, которые соответствуют группам людей с редким весом или ростом.

Тестирование сети Кохонена

Для тестирования обучения сети выделим несколько примеров, которые не участвовали в обучении:

Человек 35 лет и весом 60.

Человек 47 лет и весом 100.

Человек 52 лет и весом 77.

Отобразим их на графике синим цветом, рисунок 4.

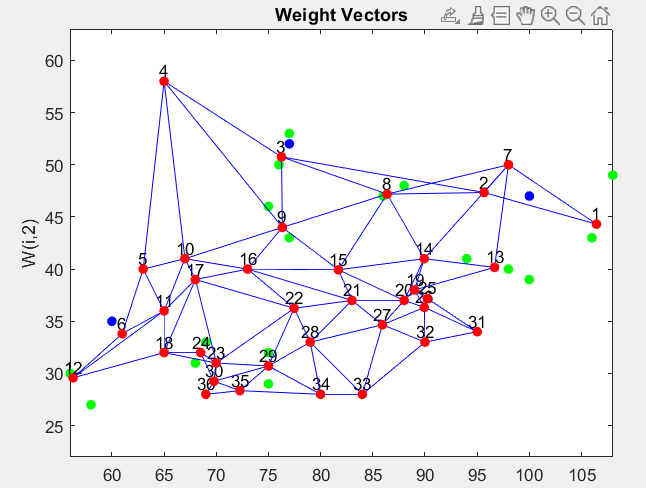


Рисунок 4 – Отображение тестовых примеров на графике синим цветом

Выполним следующий код:

res = sim(net, [60; 35])

for i = 1:length(res)

disp([num2str(i) ': ' num2str(res(i))])

end

Вывод представлен на рисунке 5.

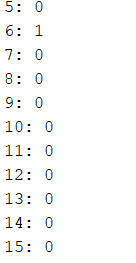


Рисунок 5 – Тестирование принадлежности человека 35 лет и весом 60 к кластеру 6.

Как видно из рисунка 5, тестирование прошло успешно – человек 35 лет и весом 60 был отнесен в группу 6, что соответствует графике, где он близок к узлу под номером 6.

Проделаем аналогичные действия для человека с 47 и весом 100, а также 52 лет и весом 77, рисунки 6 и 7.

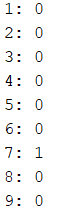


Рисунок 6 – Тестирование человека 47 лет и весом 100.

На графике видно, что человек 47 лет и 100 близок к кластерам 2, 7 и 1. Но ближе всего он оказался к узлу под номером 7 (рисунок 6), что соответствует графику на рисунке 4.

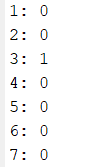


Рисунок 7 – Тестирование человека 52 лет и весом 77.

Судя по рисунку 7, человек 52 лет и весом 77 соответствует группе 3, что полностью соответствует рисунку 4.

Выводы по сети Кохонена

В данном разделе была создана сеть Кохонена в программе Matlab. Для неё была найдена обучающая выборка, состоящая из роста и веса людей. Сеть была обучена, в результате чего экземпляры выборки были отнесены к определенным кластерам, что было показано на графике. Для тестирования использовались 3 примера, которые не участвовали в обучении сети.

3.2 Сеть LVQ

Создание и обучение

Для создания сети LVQ возьмем обучающую последовательность, используя результаты сети Кохонена.

Разобьем людей на 12 классов с помощью сети Кохонена. Заменим точки числами их классов, также зададим им разный цвет следующим кодом:

T = []

for i = 1:length(indata)

y = net([indata(1, i); indata(2, i)]);

cluster\_index = vec2ind(y);

T = [T cluster\_index];

end

colors = ['r', 'r', 'r', 'r', 'b','b','b','b','g','g','g','g','c','c','c','c']

figure(1), clf, axis([min(indata(1,:))-5,max(indata(1,:))+5,min(indata(2,:))-5,max(indata(2,:)+5)]), hold on

for i = 1:12

tmp = find(T==i);

s = num2str(i) + 'r'

% plot(indata(1,tmp),indata(2,tmp), s)

text(indata(1,tmp),indata(2,tmp), num2str(i),'Color',colors(i))

hold on

end

Результат представлен на рисунке 8.

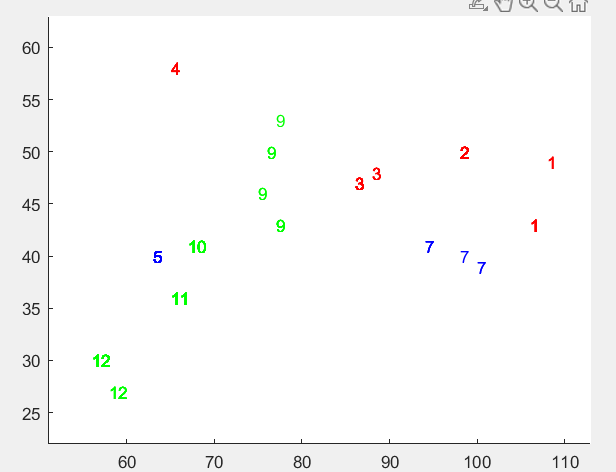


Рисунок 8 – Визуальное представление классов

Создадим сеть LVQ и обучим её:

t = ind2vec(T);

lnet = lvqnet(18,0.2,'learnlv2');

lnet.trainParam.epochs=150;

lnet=train(lnet,indata,t);

Процесс обучения представлен на рисунке 9.

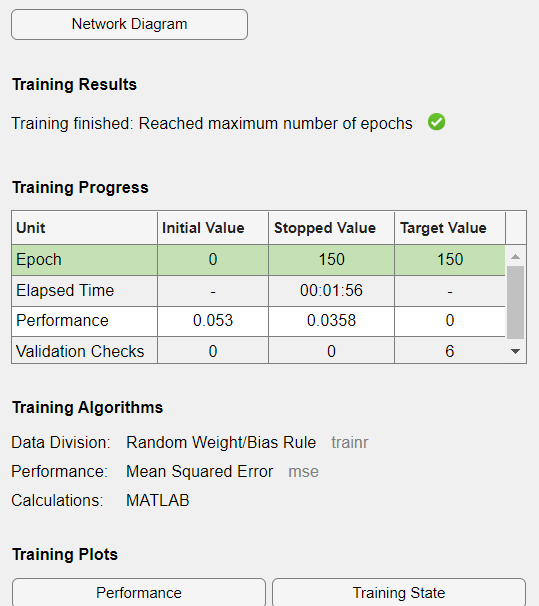


Рисунок 9 – Процесс обучения сети LVQ

После обучения получаем следующие весовые коэффициенты нейронов конкурирующего слоя, которые определяют положения центров кластеризации:

Отобразим их на графике (рисунок 10, цвет точек центров не зависит от цветов изначальных кластеров):

plotvec(lnet.IW{1}',vec2ind(lnet.LW{2}),'o');

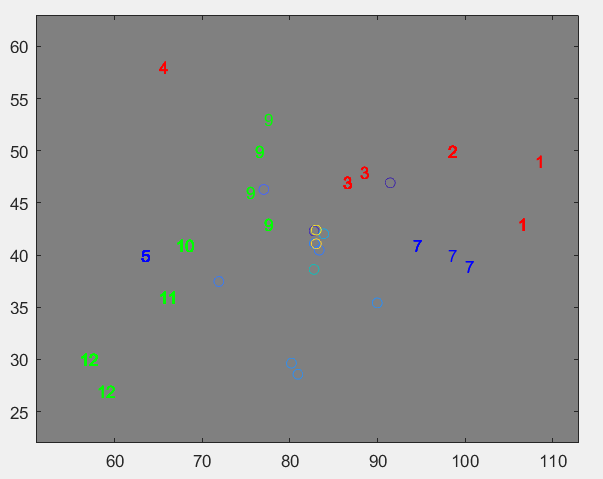


Рисунок 10 – Центры кластеризации

Наибольшее скопление кластеров произошло на стыке 1, 2 групп.

Тестирование сети

Для тестирования сети возьмем 3 примера:

Человек 35 лет и весом 60.

Человек 47 лет и весом 100.

Человек 52 лет и весом 77.

Тестирование показано на рисунке 12.

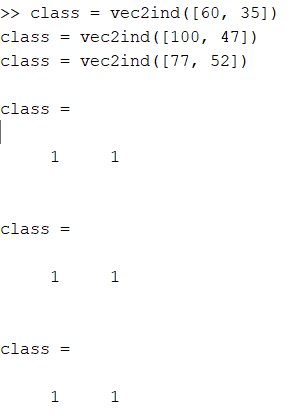


Рисунок 12 – Тестирование сети

Как видно из тестирования, есть небольшие смещения центров кластеров относительно изначальных групп, поэтому ответы немного не совпадают с их классами. Например, человек 47 лет и весом 100 по факту принадлежит 3 классу, но тестирование показало 1, но при этом он также расположен не далеко от этого класса.

Человек 35 лет и весом 60 расположен в 1 классе, поэтому он отнесен к классу под номером 1. Аналогично и с человеком 52 лет и весом 77.

Выводы по сети LVQ

В данном разделе была создана сеть LVQ в программе Matlab. В отличие от сети Кохонена, сеть LVQ требует разбиение входных векторов на классы, поэтому для обучения сети были взяты результаты обучения сети Кохонена. Сеть была обучена, в результате чего экземпляры выборки были отнесены к определенным кластерам, что было показано на графике. Для тестирования использовались 3 примера, которые не участвовали в обучении сети.

4. Выводы

В данной лабораторной работе были изучены основные команды создания, обучения и применения нейронных сетей кластеризации и классификации данных в Neural Network Toolbox с помощью сетей Кохонена и LVQ.

Приложение А

Датасет

90 33

98 50

89 38

68 39

90 33

89 38

80 28

65 36

95 34

88 37

65 36

65 36

65 36

89 38

89 38

67 41

89 38

89 38

69 33

86 47

84 28

89 38

80 28

65 36

90 33

80 28

90 33

75 29

90 33

89 38

89 38

88 48

88 37

88 48

89 38

80 28

90 33

65 32

88 48

65 36

58 27

83 37

89 38

106 43

95 34

83 37

89 38

73 40

65 36

73 40

65 36

69 28

69 28

83 37

90 33

98 50

69 28

89 38

70 31

86 47

89 38

65 36

89 38

83 37

98 50

56 30

89 38

69 28

83 37

69 28

86 47

73 40

89 38

69 28

65 36

73 40

69 28

90 33

80 28

65 36

89 38

69 28

89 38

63 40

73 40

95 34

69 33

73 40

69 28

95 34

69 28

63 40

69 28

65 36

86 47

69 28

90 33

73 40

86 47

83 37

98 50

88 37

69 28

65 36

83 37

80 28

69 28

65 36

69 28

80 28

83 37

67 41

69 28

69 28

83 37

83 37

95 34

69 28

58 27

69 28

69 28

83 37

69 28

83 37

83 37

83 37

83 37

83 37

83 37

56 30

90 33

83 37

58 27

83 37

83 37

69 28

90 33

58 27

67 41

89 38

95 34

69 33

65 36

90 33

76 50

76 50

69 28

69 28