|  |  |
| --- | --- |
| **Gerb-BMSTU_01** | **Министерство науки и высшего образования Российской Федерации**  Калужский филиал  федерального государственного бюджетного  образовательного учреждения высшего образования  ***«Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)»***  ***(КФ МГТУ им. Н.Э. Баумана)*** |

**ФАКУЛЬТЕТ** ***ИУ-КФ «Информатика и управление»***

**КАФЕДРА** \_\_***ИУ4-КФ «Программное обеспечение ЭВМ, информационные технологии»***

**ЛАБОРАТОРНАЯ РАБОТА №5**

«**Методы классификации многомерных объектов пересекающихся классов в условиях кластеризации исследуемых множеств**»

**ДИСЦИПЛИНА: «Методы машинного обучения»**

|  |  |  |
| --- | --- | --- |
| Выполнил: студент гр. ИТД.Б-61 | | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (\_\_\_\_Бурцев В.А.\_\_)  (Подпись) (Ф.И.О.) |
| Проверил: | | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (\_\_\_\_ Кручинин И.И.        )  (Подпись) (Ф.И.О.) |
| Дата сдачи (защиты):  Результаты сдачи (защиты): | | |
|  | - Балльная оценка:  - Оценка: | |

Калуга , 2020

**Цель:**

изучение методов классификации многомерных объектов пересекающихся классов в условиях кластеризации исследуемых множеств.

**Задание на лабораторную работу**

**Вариант 4**

Разработать классификатор распознавания на основе параметров глазного яблока с использованием алгоритмов К-средних, ЕМ, РАМ для идентификации 42 сотрудников оборонного предприятия для доступа в зону В средней секретности. Всего применим 10 параметров – характеристик: роговица, радужка, лимб, конъюнктива, хрусталик, сетчатка, хориодея, склера, гиалоидная мембрана, венозный синус. Для достижения положительного результата достаточно 70 % совпадения характеристик.

Для алгоритма ЕМ в функции ЕМ использовать параметр ModelName = “EEE”, параметр parameters variance = “Cholsigma”

Для алгоритма РАМ использовать параметр metric=’euclidean’, параметр medoids= 5:1

Для алгоритма Kmeans максимальное число кластеров = 21, параметр iter.max=19, параметр algorithm= “MacQueen”, centers = 23, num.seeds=27

Результаты классификации трех алгоритмов сравнить с точки зрения уровня ошибки распознавания и представить в графическом виде.

**Ход выполнения работы**

Листинг программы на R

rm(list = ls()) #очистка памяти

library(cluster)

#library(factoextra)

ramFo = data.frame(read.table("C:\\Василий\\WORK\\КФ МГТУ им. Н.Э.Баумана\\Бакалавриат\\3 курс\\Бурцев В.А.ИТД.Б-61\\Методы машинного обучения\\Лабораторные работы\\Лабораторная работа№5\\APP005\_1.txt", header = TRUE, sep = ""))

print("Исходные данные")

print(ramFo)

df.stand <- as.data.frame(scale(ramFo[2:11]))

set.seed(27)

c(kmeans(df.stand, centers = 23, iter.max = 19, algorithm= "MacQueen",nstart = 1)$tot.withinss,

kmeans(df.stand, centers = 23, iter.max = 19, algorithm= "MacQueen",nstart = 27)$tot.withinss)

k.max <- 21 # максимальное число кластеров

wss <- sapply(1:k.max, function(k){

kmeans(df.stand, k,nstart = 27)$tot.withinss

})

plot(1:k.max, wss, type = "b", pch = 19, frame = FALSE,

xlab = "Число кластеров K",

ylab = "Общая внутригрупповая сумма квадратов")

gap\_stat <- clusGap(df.stand, FUN = pam, K.max = 21, B = 100)

print(gap\_stat, method = "firstSEmax")

(k.pam <- pam(df.stand, k = 5, metric = "euclidean", medoids = 5:1))

library(mclust)

msEst <- mstep(modelName = "EEE", data = ramFo[,2:11],

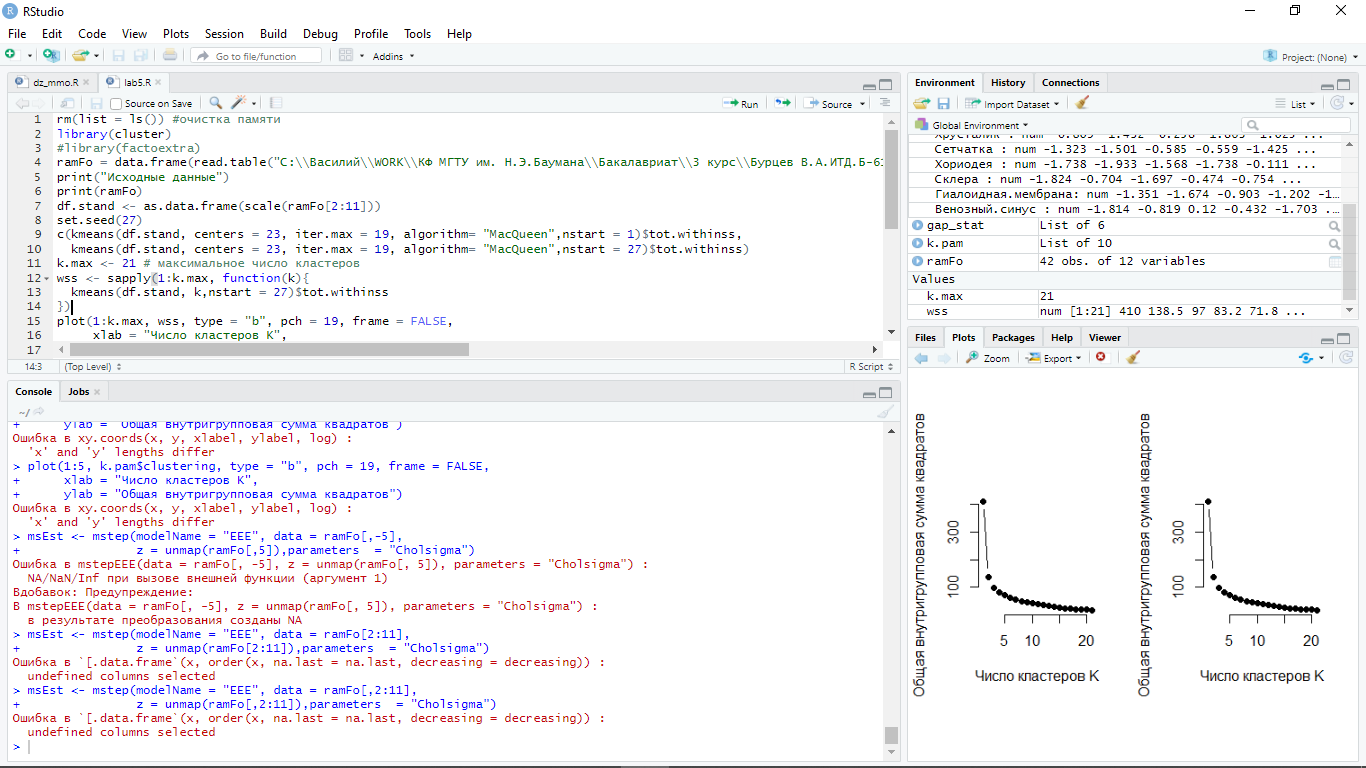
z = unmap(ramFo[,2:11]),parameters = "Cholsigma")

names(msEst)

em(modelName = msEst$modelName, data = ramFo[2:11],

parameters = "Cholsigma")

Результат программы



**Вывод:**

в ходе лабораторной работы были изучены методы классификации многомерных объектов пересекающихся классов в условиях кластеризации исследуемых множеств.