|  |  |
| --- | --- |
| **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 №2**

## «Метрические методы классификации многомерных объектов пересекающихся классов»

**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:**

сформировать практические навыки разработки программ с использованием метрических методов классификации многомерных объектов пересекающихся классов на языке R.

**Задание на лабораторную работу**

**Вариант 4**

1. Определим уровень финансовой устойчивости предприятия, как составной части общей устойчивости предприятия (при этом соблюдаются сбалансированность финансовых потоков, наличие средств, позволяющих организации поддерживать свою деятельность в течение определенного периода времени, в том числе обслуживая полученные кредиты и производя продукцию).

Исходные данные следует организовать в виде таблицы:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| RES | K1 | K2 | K3 | K4 | K5 | … | KN |
| T | 0.252 | 0.301 |  |  |  |  | 0.539 |
| F | 0.327 | 0.240 |  | 0.671 |  |  |  |
| … | … | … | … | … | … | … | … |
| T | 0.458 |  |  | 0.683 |  |  |  |

В первом столбце заносится значение бинарной классификации – финансовой состояние предприятия - устойчиво или нет (True, False). В данном варианте признаками финансовой устойчивости будут: Коэффициент мобильности имущества, Коэффициент мобильности оборотных средств, Коэффициент обеспеченности запасов, Коэффициент краткосрочной задолженности. Строк в таблице должно быть 120 (каждая строка - сведения по проверенному предприятию).

1. Используем метод К-ближайших соседей и метод Парзена. Сформировать обучающие и тестовые выборки. Полученные результаты визуализировать и сравнить. Представить значения параметров с минимальным уровнем ошибки. Для метода К соседей параметр К = 22, для метода Парзена тип ядра выбрать "rectangular", "triangular" а параметр optim. method = "Brent", "Nelder-Mead". Проверить точность прогнозов.

**Ход выполнения работы**

1. Определим уровень финансовой устойчивости предприятия, как составной части общей устойчивости предприятия (при этом соблюдаются сбалансированность финансовых потоков, наличие средств, позволяющих организации поддерживать свою деятельность в течение определенного периода времени, в том числе обслуживая полученные кредиты и производя продукцию).

Исходные данные следует организовать в виде таблицы:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| RES | K1 | K2 | K3 | K4 | K5 | … | KN |
| T | 0.252 | 0.301 |  |  |  |  | 0.539 |
| F | 0.327 | 0.240 |  | 0.671 |  |  |  |
| … | … | … | … | … | … | … | … |
| T | 0.458 |  |  | 0.683 |  |  |  |

В первом столбце заносится значение бинарной классификации – финансовой состояние предприятия - устойчиво или нет (True, False). В данном варианте признаками финансовой устойчивости будут: Коэффициент мобильности имущества, Коэффициент мобильности оборотных средств, Коэффициент обеспеченности запасов, Коэффициент краткосрочной задолженности. Строк в таблице должно быть 120 (каждая строка - сведения по проверенному предприятию).

Листинг программы на R

#подготовка данных для фрейма

N = 120

RES <- c(T,F)

Kmp <- round(runif(N, 200, 900)) \* 10^-3

Kmca <- round(runif(N, 200, 900)) \* 10^-3

Rr <- round(runif(N, 200, 900)) \* 10^-3

Rstd <- round(runif(N, 200, 900)) \* 10^-3

frame = data.frame(RES, Kmp, Kmca, Rr, Rstd)

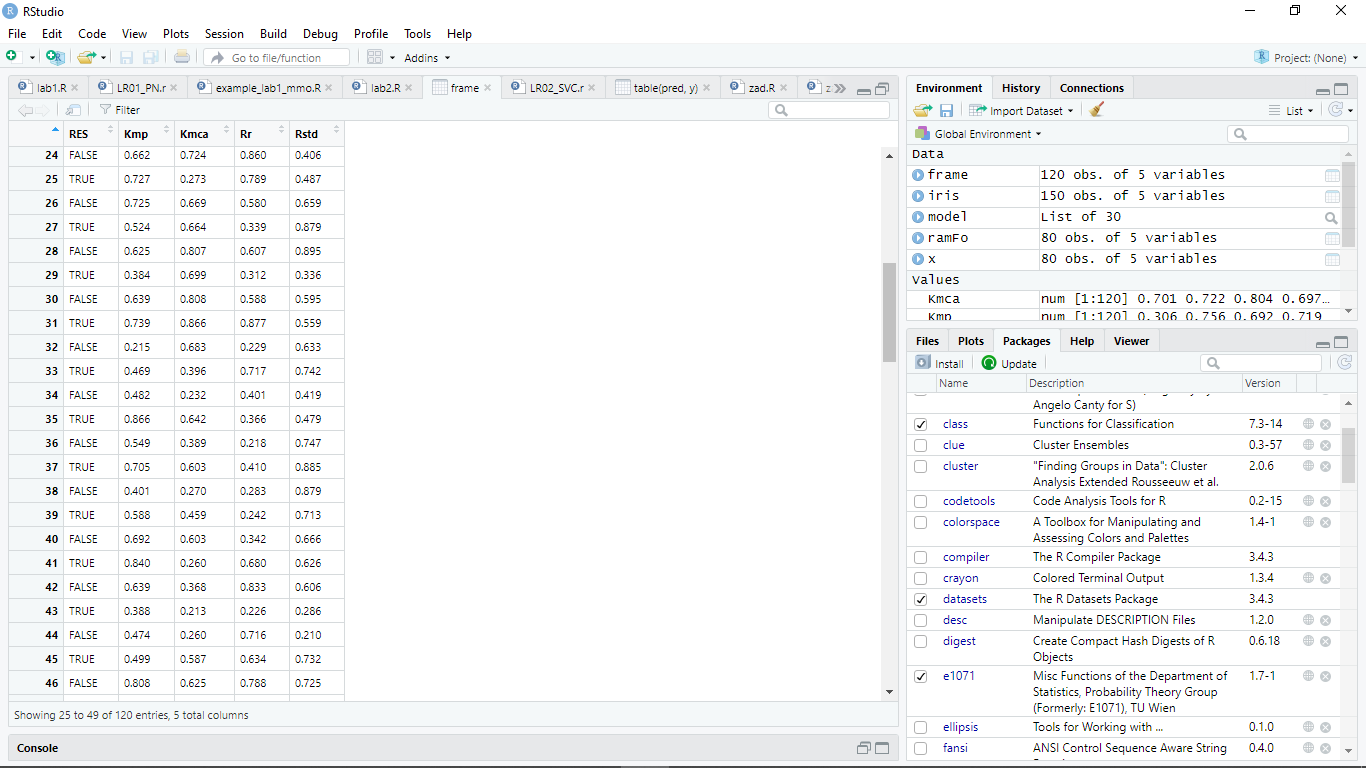
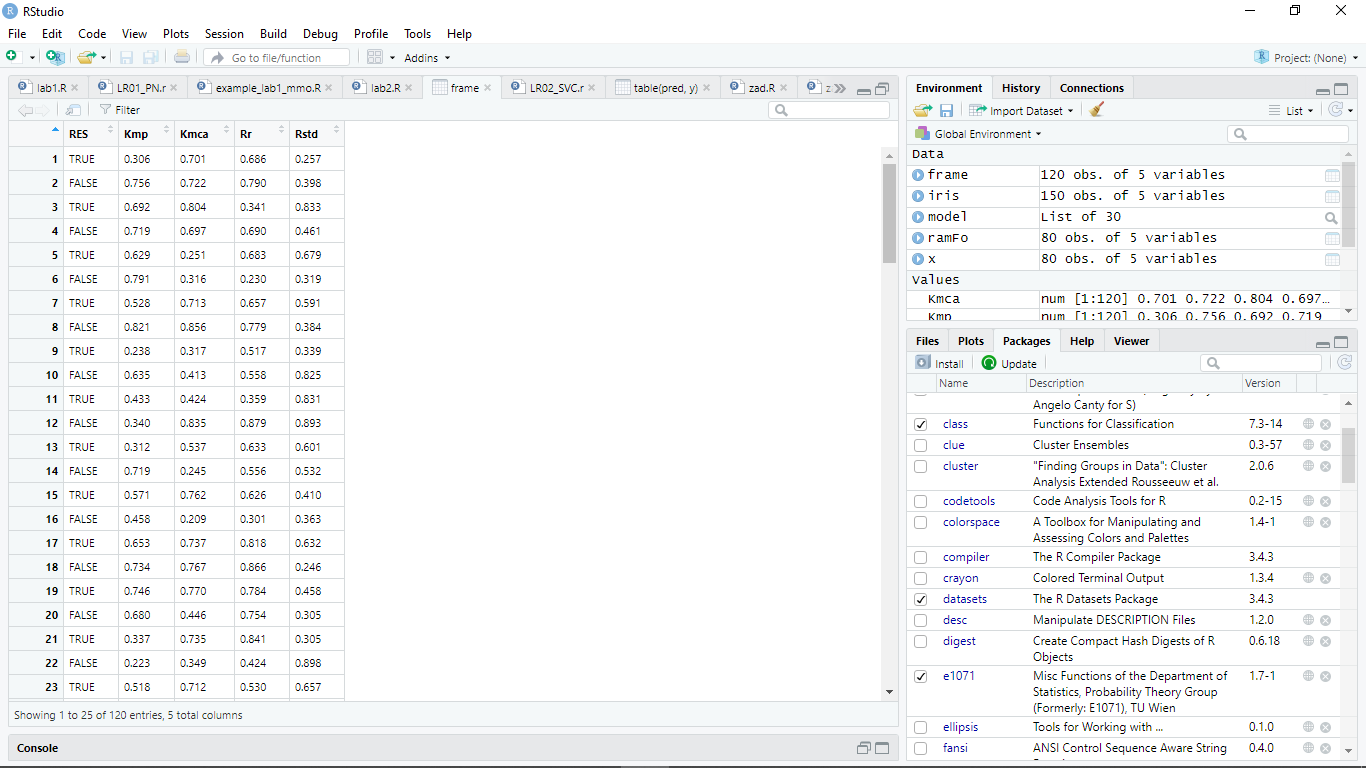
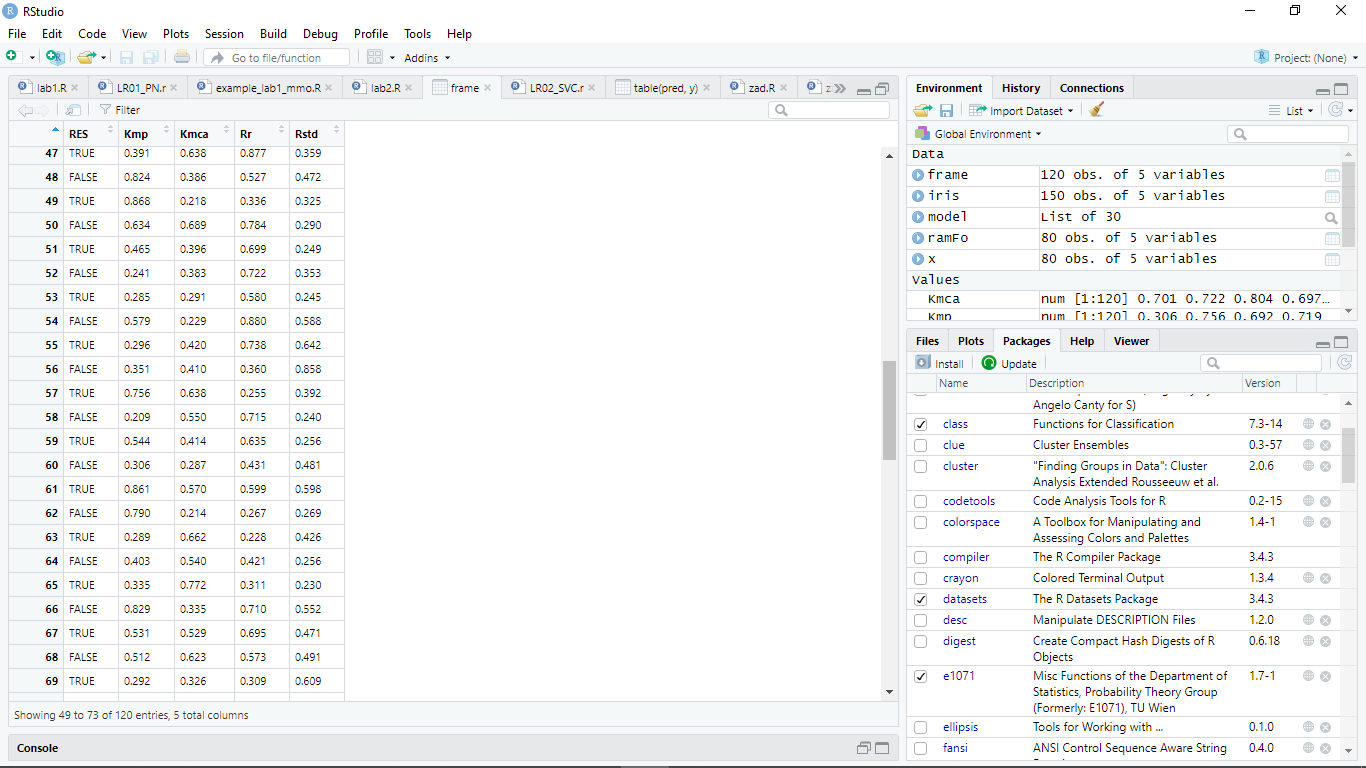
print(frame)

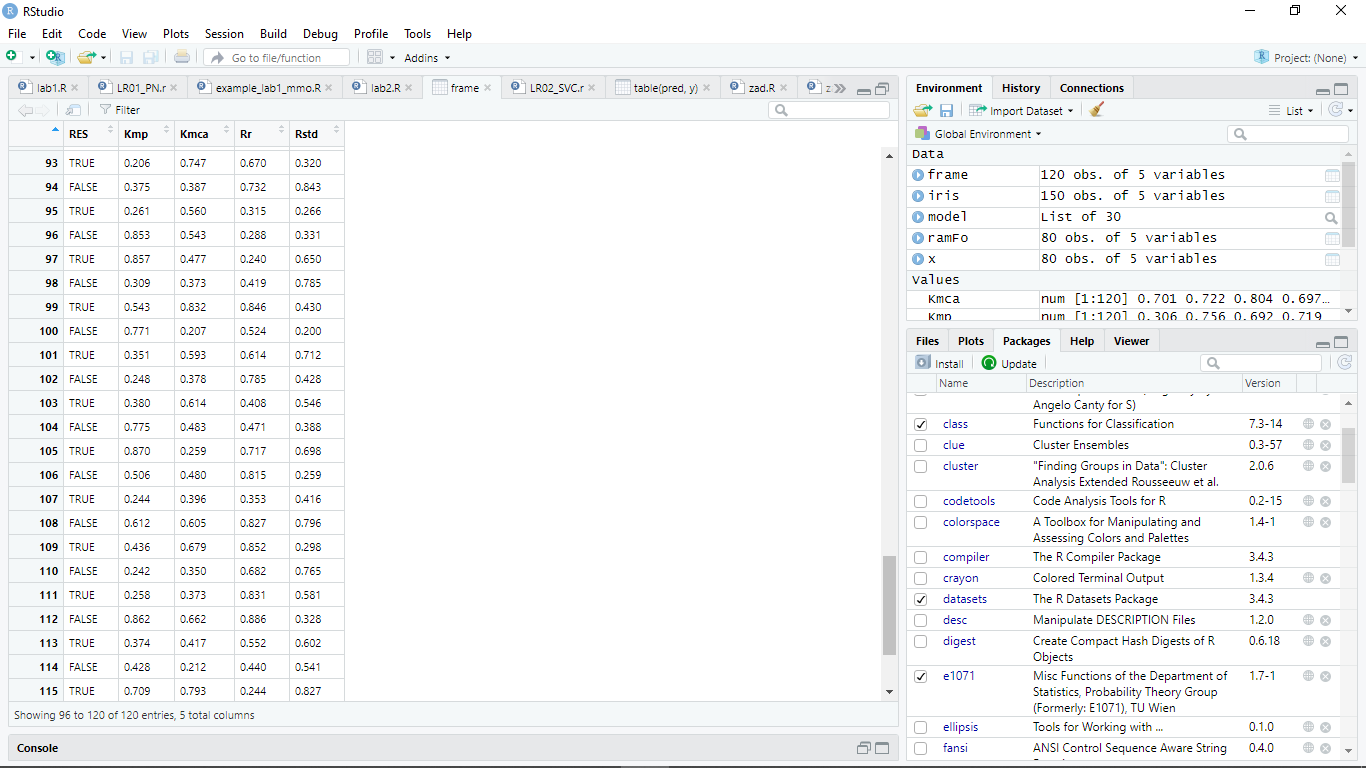
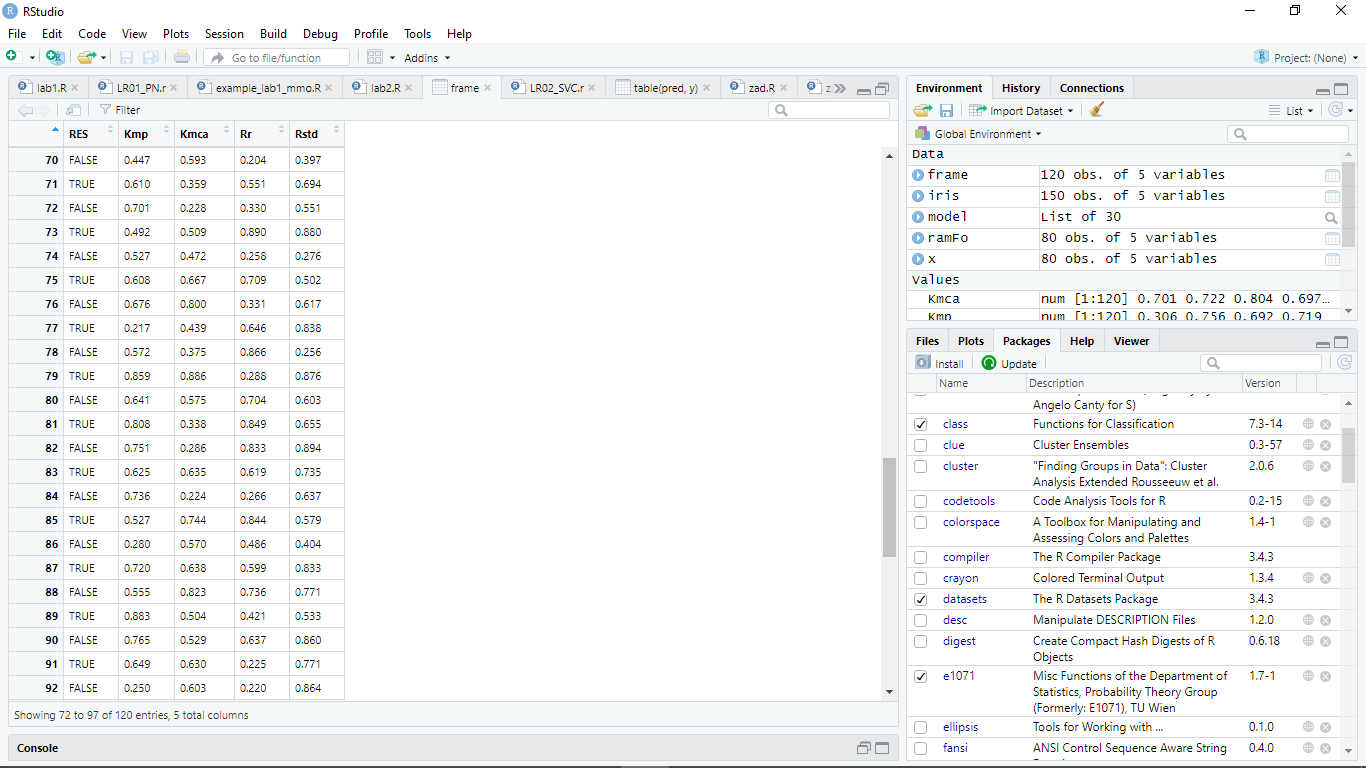
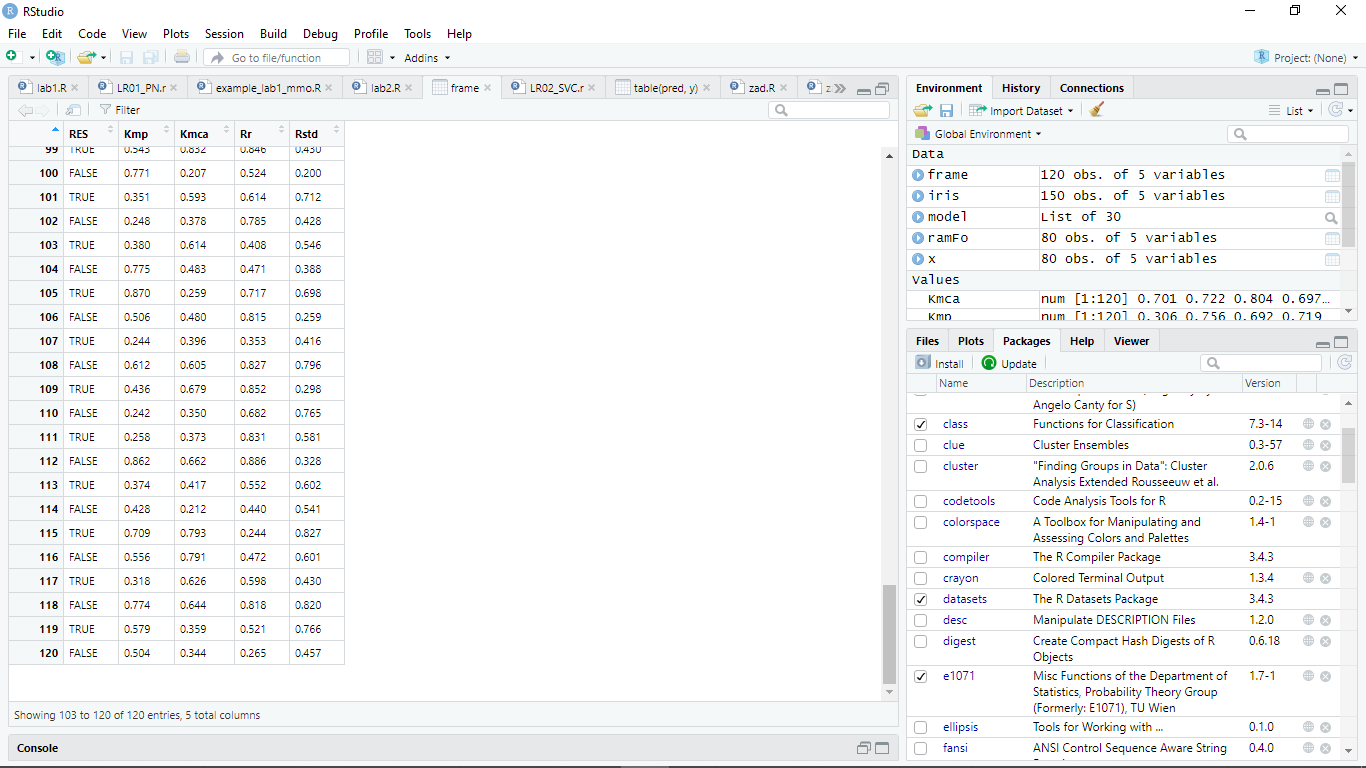
#созданиефрейма

frame = data.frame(RES, Kmp, Kmca, Rr, Rstd)

print(frame)

Результат программы – созданный фрейм

1. Используем метод К-ближайших соседей и метод Парзена. Сформировать обучающие и тестовые выборки. Полученные результаты визуализировать и сравнить. Представить значения параметров с минимальным уровнем ошибки. Для метода К соседей параметр К = 22, для метода Парзена тип ядра выбрать "rectangular", "triangular" а параметр optim. method = "Brent", "Nelder-Mead". Проверить точность прогнозов.

Листинг программы на R

А) метод K-ближайших соседей

write.table(frame, file="GMB1.txt")

ramFo2 = data.frame(read.csv("GMB1.txt", stringsAsFactors = FALSE,header = TRUE, sep = ""))

table(ramFo2$RES)

ramFo2$RES <- factor(ramFo2$RES, levels = c(TRUE, FALSE),

labels = c("Stable", "Unstable"))

round(prop.table(table(ramFo2$RES))\*100, digits = 1)

normalize <- function(x) {return ((x - min(x)) / (max(x) - min(x)))}

rm\_n <- as.data.frame(lapply(ramFo2[2:5], normalize))

rm\_train <- rm\_n[1:60, ]

rm\_test <- rm\_n[61:120, ]

rm\_train\_labels <- ramFo2[1:60, 1]

rm\_test\_labels <- ramFo2[61:120, 1]

library("class")

rm\_test\_pred <- knn(train = rm\_train, test = rm\_test,cl =

rm\_train\_labels, k=22)

library("gmodels")

CrossTable(x = rm\_test\_labels, y = rm\_test\_pred, prop.chisq=FALSE)

Результат программы

Total Observations in Table: 60

| rm\_test\_pred

rm\_test\_labels | Stable | Unstable | Row Total |

---------------|-----------|-----------|-----------|

Stable | 21 | 8 | 29 |

| 0.724 | 0.276 | 0.483 |

| 0.500 | 0.444 | |

| 0.350 | 0.133 | |

---------------|-----------|-----------|-----------|

Unstable | 21 | 10 | 31 |

| 0.677 | 0.323 | 0.517 |

| 0.500 | 0.556 | |

| 0.350 | 0.167 | |

---------------|-----------|-----------|-----------|

Column Total | 42 | 18 | 60 |

| 0.700 | 0.300 | |

---------------|-----------|-----------|-----------|

**Вывод:**

в ходе выполнения лабораторной работы были приобретены практические навыки разработки программ с использованием метрических методов классификации многомерных объектов пересекающихся классов на языке R.