



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 ИУК «Информатика и управление»

КАФЕДРА ИУК4 «Программное обеспечение ЭВМ, информационные технологии»

ЛАБОРАТОРНАЯ РАБОТА №3

«Классификация объектов по Байесу»

ДИСЦИПЛИНА: «Методы машинного обучения»

Выполнил: студент гр. ИУК4-62Б _____ (Калашников А.С.)
(Подпись) (Ф.И.О.)

Проверил: _____ (Кручинин И.И.)
(Подпись) (Ф.И.О.)

Дата сдачи (защиты):

Результаты сдачи (защиты):

- Балльная оценка:

- Оценка:

Калуга, 2023

Цель: изучение основных процедур дискриминантного анализа: дискриминации и классификации, построение и определение количества дискриминантных функций и их разделительной способности, нахождение классифицирующих функций с использованием функций Фишера и расстояния Махаланобиса.

Задание:

- среднее значение переменных внутри классов, общее среднее;
- матрицу перекрестных произведений и ковариационную матрицу общего рассеяния;
- матрицу внутригрупповых квадратов и перекрестных произведений и корреляционную матрицу;
- матрицу межгрупповых квадратов и перекрестных произведений и корреляционную матрицу;
- коэффициенты канонической дискриминантной функции;
- коэффициенты классифицирующей функции Фишера;
- используя оценки априорных вероятностей принадлежности объектов к группам, определить расстояние Махаланобиса;
- вычислить обобщенное расстояние Рао и его значимость.

Листинг:

```
Out_CTab<- function(model, group, type = "lda") {
  classified<- predict(model)$class
  t1 <- table(group, classified)
  Err_S<- mean(group != classified)
  mahDist<- NA
  if (type == "lda") {
    mahDist<- dist(model$means %*% model$scaling)
  }
  t2 <- table(group, update(model, CV = T)$class -> LDA.cv)
  Err_CV<- mean(group != LDA.cv)
  Err_S.MahD<- c(Err_S, mahDist)
  Err_CV.N<- c(Err_CV, length(group))
  cbind(t1, Err_S.MahD, t2, Err_CV.N)
}

library(MASS)
library(klaR)
library(biotools)

Diseases <- read.table(file = "Diseases.txt", sep = ",", header =
TRUE, row.names = 1)
lda.all<- lda(FAC ~ ., data = Diseases[, -10])
Out_CTab(lda.all, Diseases$FAC)
stepclass(FAC ~ ., data = Diseases[, -10], method = "lda")
lda.step<- lda(FAC ~ Hemog + Bi, data = Diseases[, -10])
```

```

ldaProfile<- rfe(Diseases[, 1:9], Diseases$FAC, sizes = 2:9,rfeControl
= rfeControl(functions = ldaFuncs, method = "repeatedcv", repeats =
6))
Diseases$FAC<- as.factor(ifelse(Diseases$Class == 3, "C2", "C1"))

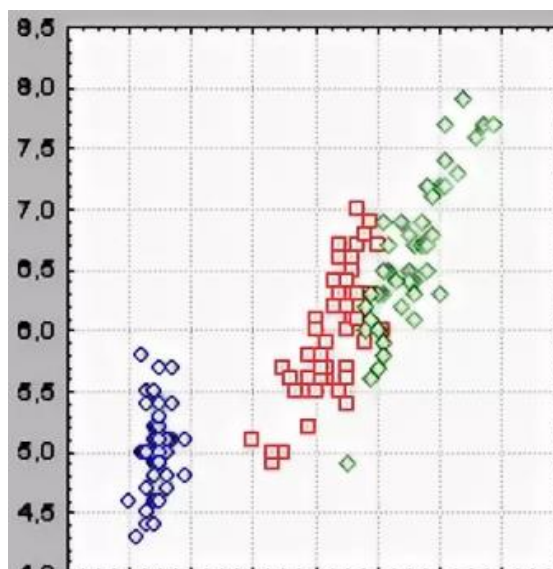
lda.full.pro<- train(Diseases[, 1:9], Diseases$FAC,data = Diseases,
method = "lda",
trControl = trainControl(method = "repeatedcv", repeats = 5,
classProbs = TRUE), metric = "Accuracy")

lda.step.pro <- train(FAC ~ A1 + Hemat, data = Diseases, method =
"lda",
trControl = trainControl(method = "repeatedcv", repeats = 5,
classProbs = TRUE), metric = "Accuracy")

lda.rfe.pro <- train(FAC ~ A1 + Hemog + Bi, data = Diseases, method =
"lda",
trControl = trainControl(method = "repeatedcv", repeats = 5,
classProbs = TRUE), metric = "Accuracy")
plot(varImp(lda.full.pro))

```

Результаты работы:



Вывод: в ходе выполнения лабораторной работы были изучены основные процедуры дискриминантного анализа: дискриминации и классификации, построение и определение количества дискриминантных функций и их разделительной способности, нахождение классифицирующих функций с использованием функций Фишера и расстояния Махаланобиса.