



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. ИУК4-62Б \_\_\_\_\_ ( \_\_\_\_\_ Борисов Н.С. )  
(Подпись) (Ф.И.О.)

Проверил: \_\_\_\_\_ ( \_\_\_\_\_ Кручинин И.И. )  
(Подпись) (Ф.И.О.)

Дата сдачи (защиты):

Результаты сдачи (защиты):

- Балльная оценка:

- Оценка:

Калуга, 2022

**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).

2. Используем метод K-ближайших соседей и метод Парзена. Сформировать обучающие и тестовые выборки. Полученные результаты визуализировать и сравнить. Представить значения параметров с минимальным уровнем ошибки. Для метода K соседей параметр K = 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)
#создание фрейма
frame = data.frame(RES, Kmp, Kmca, Rr, Rstd)
print(frame)
```

## Результат программы – созданный фрейм

	RES	Kmp	Kmca	Rr	Rstd												
1	TRUE	0.306	0.701	0.686	0.257	24	FALSE	0.662	0.724	0.860	0.406	47	TRUE	0.391	0.638	0.877	0.359
2	FALSE	0.756	0.722	0.790	0.398	25	TRUE	0.727	0.273	0.789	0.487	48	FALSE	0.824	0.386	0.527	0.472
3	TRUE	0.692	0.804	0.341	0.833	26	FALSE	0.725	0.669	0.580	0.659	49	TRUE	0.868	0.218	0.336	0.325
4	FALSE	0.719	0.697	0.690	0.461	27	TRUE	0.524	0.664	0.339	0.879	50	FALSE	0.634	0.689	0.784	0.249
5	TRUE	0.629	0.251	0.663	0.679	28	FALSE	0.625	0.807	0.607	0.895	51	TRUE	0.465	0.396	0.699	0.290
6	FALSE	0.791	0.316	0.230	0.319	29	TRUE	0.384	0.699	0.312	0.336	52	FALSE	0.241	0.383	0.722	0.353
7	TRUE	0.528	0.713	0.657	0.591	30	FALSE	0.639	0.808	0.588	0.595	53	TRUE	0.285	0.291	0.580	0.245
8	FALSE	0.821	0.856	0.779	0.384	31	TRUE	0.739	0.866	0.877	0.559	54	FALSE	0.579	0.229	0.880	0.588
9	TRUE	0.238	0.317	0.517	0.339	32	FALSE	0.215	0.683	0.229	0.633	55	TRUE	0.296	0.420	0.738	0.642
10	FALSE	0.635	0.413	0.558	0.825	33	TRUE	0.469	0.396	0.717	0.742	56	FALSE	0.351	0.410	0.360	0.858
11	TRUE	0.433	0.424	0.359	0.831	34	FALSE	0.482	0.232	0.401	0.419	57	TRUE	0.756	0.638	0.255	0.392
12	FALSE	0.340	0.835	0.879	0.893	35	TRUE	0.866	0.642	0.366	0.479	58	FALSE	0.209	0.550	0.715	0.240
13	TRUE	0.312	0.537	0.633	0.601	36	FALSE	0.549	0.389	0.218	0.747	59	TRUE	0.544	0.414	0.635	0.256
14	FALSE	0.719	0.245	0.556	0.532	37	TRUE	0.705	0.603	0.410	0.885	60	FALSE	0.306	0.287	0.431	0.481
15	TRUE	0.571	0.762	0.626	0.410	38	FALSE	0.401	0.270	0.283	0.879	61	TRUE	0.861	0.570	0.599	0.598
16	FALSE	0.458	0.209	0.301	0.363	39	TRUE	0.588	0.459	0.242	0.713	62	FALSE	0.790	0.214	0.267	0.269
17	TRUE	0.653	0.737	0.818	0.632	40	FALSE	0.692	0.603	0.342	0.666	63	TRUE	0.289	0.662	0.228	0.426
18	FALSE	0.734	0.767	0.866	0.246	41	TRUE	0.840	0.260	0.680	0.626	64	FALSE	0.403	0.540	0.421	0.556
19	TRUE	0.746	0.770	0.784	0.458	42	FALSE	0.639	0.368	0.833	0.606	65	TRUE	0.335	0.772	0.311	0.230
20	FALSE	0.680	0.446	0.754	0.305	43	TRUE	0.388	0.213	0.226	0.286	66	FALSE	0.829	0.335	0.710	0.552
21	TRUE	0.337	0.735	0.841	0.305	44	FALSE	0.474	0.260	0.716	0.210	67	TRUE	0.531	0.529	0.695	0.471
22	FALSE	0.223	0.349	0.424	0.898	45	TRUE	0.499	0.587	0.634	0.732	68	FALSE	0.512	0.623	0.573	0.491
23	TRUE	0.518	0.712	0.530	0.657	46	FALSE	0.808	0.625	0.788	0.725	69	TRUE	0.292	0.326	0.309	0.609

70	FALSE	0.447	0.593	0.204	0.397
71	TRUE	0.610	0.359	0.551	0.694
72	FALSE	0.701	0.228	0.330	0.551
73	TRUE	0.492	0.509	0.890	0.880
74	FALSE	0.527	0.472	0.258	0.276
75	TRUE	0.608	0.667	0.709	0.502
76	FALSE	0.676	0.800	0.331	0.617
77	TRUE	0.217	0.439	0.646	0.838
78	FALSE	0.572	0.375	0.866	0.256
79	TRUE	0.859	0.886	0.288	0.876
80	FALSE	0.641	0.575	0.704	0.603
81	TRUE	0.808	0.338	0.849	0.655
82	FALSE	0.751	0.286	0.833	0.894
83	TRUE	0.625	0.635	0.619	0.735
84	FALSE	0.736	0.224	0.266	0.637
85	TRUE	0.527	0.744	0.844	0.579
86	FALSE	0.280	0.570	0.486	0.404
87	TRUE	0.720	0.638	0.599	0.833
88	FALSE	0.555	0.823	0.736	0.771
89	TRUE	0.883	0.504	0.421	0.533
90	FALSE	0.765	0.529	0.637	0.860
91	TRUE	0.649	0.630	0.225	0.771
92	FALSE	0.250	0.603	0.220	0.864
93	TRUE	0.206	0.747	0.670	0.320
94	FALSE	0.375	0.387	0.732	0.843
95	TRUE	0.261	0.560	0.315	0.266
96	FALSE	0.853	0.543	0.288	0.331
97	TRUE	0.857	0.477	0.240	0.650
98	FALSE	0.309	0.373	0.419	0.785
99	TRUE	0.543	0.832	0.846	0.430
100	FALSE	0.771	0.207	0.524	0.200
101	TRUE	0.351	0.593	0.614	0.712
102	FALSE	0.248	0.378	0.785	0.428
103	TRUE	0.380	0.614	0.408	0.546
104	FALSE	0.775	0.483	0.471	0.388
105	TRUE	0.870	0.259	0.717	0.698
106	FALSE	0.506	0.480	0.815	0.259
107	TRUE	0.244	0.396	0.353	0.416
108	FALSE	0.612	0.605	0.827	0.796
109	TRUE	0.436	0.679	0.852	0.298
110	FALSE	0.242	0.350	0.682	0.765
111	TRUE	0.258	0.373	0.831	0.581
112	FALSE	0.862	0.662	0.886	0.328
113	TRUE	0.374	0.417	0.552	0.602
114	FALSE	0.428	0.212	0.440	0.541
115	TRUE	0.709	0.793	0.244	0.827
116	FALSE	0.556	0.791	0.472	0.601
117	TRUE	0.318	0.626	0.598	0.430
118	FALSE	0.774	0.644	0.818	0.820
119	TRUE	0.579	0.359	0.521	0.766
120	FALSE	0.504	0.344	0.265	0.457

- 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 K = 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 К-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.