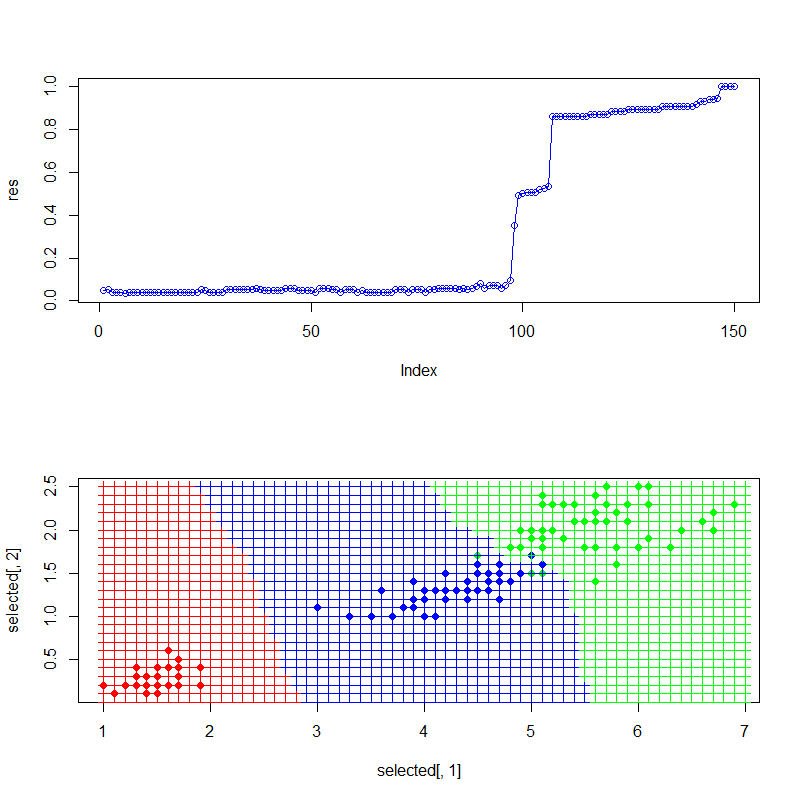
Тема: LOO для kNN. Балицкая Анастасия 401-И



Вывод: Минимальный LOO достигается при k=6

Код программы:

dist **=** **function(**u, v**)** **{** #это возвращает эвклидовое расстояние между двумя объектами

features **=** length**(**u**)-**1

sqrt**(**sum**((**u**[**1**:**features**]-**v**[**1**:**features**])^**2**))**

**}**

distances **=** **function(**obj, data, metric**)** **{** # это возвращает отсортированный набор данных по метрике для объекта

cases **=** dim**(**data**)[**1**]**

dists **=** matrix**(**0, cases, 2**)**

**for** **(**i **in** 1**:**cases**)** **{**

cost **=** metric**(**obj, data**[**i,**])**

dists**[**i,**]** **=** c**(**cost, i**)**

**}**

idx **=** order**(**dists**[**,1**])**

data**[**dists**[**idx,2**]**,**]**

**}**

kNN **=** **function(**obj, data, k, metric**=**dist**)** **{** # это k-ближайших соседей

sorted **=** distances**(**obj, data, metric**)**

applykNN**(**sorted, k**)**

**}**

applykNN **=** **function(**sorted, k**)** **{**

n **=** dim**(**sorted**)[**2**]-**1

t **=** table**(**sorted**[**1**:**k,n**+**1**])**

which.max**(**t**)**

**}**

getLoo **=** **function(**x**)** **{**

l **=** dim**(**x**)[**1**]**

n **=** dim**(**x**)[**2**]** **-** 1

maxk **=** l

loo **=** rep**(**0, times**=**maxk**)**

**for** **(**i **in** 1**:**l**)** **{**

dists **=** distances**(**x**[**i,**]**, x**[-**i,**]**, dist**)**

**for** **(**k **in** 1**:**maxk**)** **{**

class **=** applykNN**(**dists, k**)**

**if** **(**as.integer**(**class**)** **!=** as.integer**(**x**[**i,n**+**1**]))** **{**

loo**[**k**]** **=** loo**[**k**]** **+** 1

**}**

**}**

print**(**i**)**

print**(**loo**)**

**}**

loo **=** loo **/** l

return**(**loo**)**

**}**

which.min**(**res**)**

# res = getLoo(selected)

par**(**mfrow**=**c**(**2,1**))**

colors **=** c**(**"red", "blue", "green"**)**

plot**(**res, col**=**"blue"**)**

lines**(**res, col**=**"blue"**)**

plot**(**selected**[**,1**]**, selected**[**,2**]**, col**=**colors**[**selected**[**,3**]]**, pch**=**19**)**

xx **=** seq**(**from**=**1, to**=**7, by**=**0.1**)**

yy **=** seq**(**from**=**0, to**=**2.5, by**=**0.1**)**

**for** **(**x **in** xx**)** **{**

**for** **(**y **in** yy**)** **{**

obj **=** c**(**x, y, 0.0**)**

class **=** kNN**(**obj, selected, 6**)**

points**(**x, y, col**=**colors**[**class**]**, pch**=**3**)**

**}**

**}**