

Machine Learning 2019

袁欣

2019 年 5 月 28 日

1 朴素贝叶斯

1.1 朴素贝叶斯实验数据

读入朴素贝叶斯实验数据，并进行展示。

```
da <- read.csv("data.csv")
da.train <- da[1:14, ]
da.test <- da[15, ]
```

- 数据展示

No	Age	Income	Student	Credit_rating	Class.buys_computer
1	<=30	High	No	Fair	No
2	<=30	High	No	Excellent	No
3	31-40	High	No	Fair	Yes
4	>40	Medium	No	Fair	Yes
5	>40	Low	Yes	Fair	Yes
6	>40	Low	Yes	Excellent	No
7	31-40	Low	Yes	Excellent	Yes
8	<=30	Medium	No	Fair	No
9	<=30	Low	Yes	Fair	Yes
10	>40	Medium	Yes	Fair	Yes
11	<=30	Medium	Yes	Excellent	Yes
12	31-40	Medium	No	Excellent	Yes
13	31-40	High	Yes	Fair	Yes
14	>40	Medium	No	Excellent	No
15	<=30	Medium	Yes	Fair	

1.2 朴素贝叶斯简介

朴素贝叶斯分类器 (naive Bayes classifier) 采用了“属性条件独立性假设”：对已知类别，假设所有属性相互独立。换言之，假设每个属性独立地对分裂结果发生影响。

- 代码实现

Classn	P
No	0.0004036
Yes	0.0014952

由上表可以看出 Yes 的概率大于 No 的概率，所以测试样本应该会买电脑。