bayes

提前将训练集的数据写进了一个csv的文件，然后用一个函数就可以把他们读进来。

train\_data <- read.csv('buy',sep = ',')  
nrow(train\_data)

## [1] 14

buy\_computer <- train\_data[train\_data$buys\_computer=='yes',]  
not\_buy\_computer <- train\_data[train\_data$buys\_computer=='no',]

buyer <- data.frame(age = '<30',imcome = 'medium',student = 'yes',credit\_rating= 'fair')

#### 面向过程的编程计算。

1.计算全体购买的概率：p(b) 2.计算全体不购买的概率：p(nb)

p\_b <- nrow(buy\_computer)/nrow(train\_data)  
p\_nb <- nrow(not\_buy\_computer)/nrow(train\_data)

3.计算在buy\_computer 和not\_buy\_computer的顾客中age<30，imcome=medium,student=yes,credit\_rating=fair的概率

p\_a <- sum(buy\_computer$age=='<30')/nrow(buy\_computer)  
p\_na <- sum(not\_buy\_computer=='<30')/nrow(not\_buy\_computer)  
p\_i <- sum(buy\_computer$income=='medium')/nrow(buy\_computer)  
p\_ni <- sum(not\_buy\_computer$income=='medium')/nrow(not\_buy\_computer)  
p\_s <- sum(buy\_computer$student =='yes')/nrow(buy\_computer)  
p\_ns <- sum(not\_buy\_computer$student == 'yes')/nrow(not\_buy\_computer)  
p\_c <- sum(buy\_computer$credit\_rating =='fair')/nrow(buy\_computer)  
p\_nc <- sum(not\_buy\_computer$credit\_rating=='fair')/nrow(not\_buy\_computer)

4.区分这个顾客究竟是属于哪个类别的。

py <- p\_b\*p\_a\*p\_i\*p\_s\*p\_c  
pn <- p\_nb\*p\_na\*p\_ni\*p\_ns\*p\_nc

if(py>=pn){  
 print('this cumstomer will buy computer')  
}else{  
 print('this cumstorm will not buy computer')  
}

## [1] "this cumstomer will buy computer"