Machine Learning 2019

袁欣

2019 年 4 月 28 日

1 神经网络

1.1 西瓜数据 3.0

读入西瓜数据 3.0，对分类变量进行独热编码（One-hot encoding）。

- 数据展示

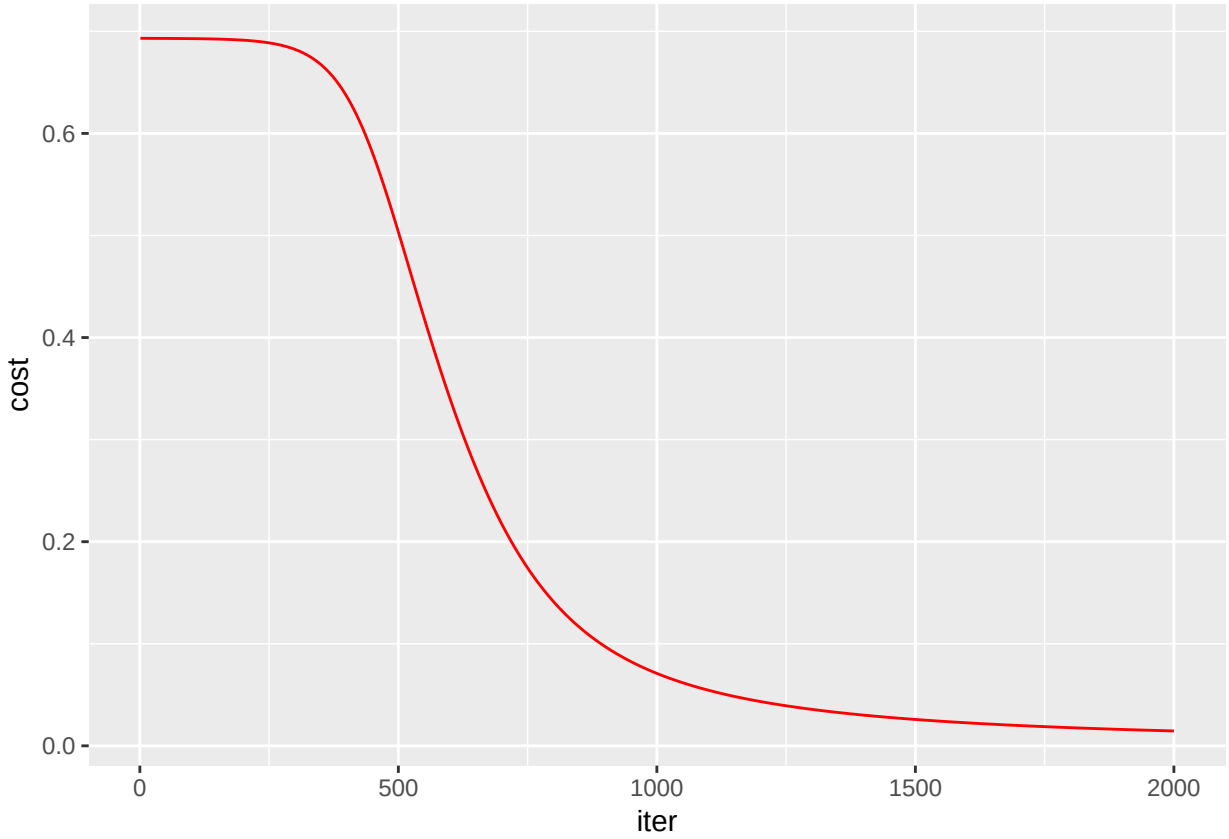
色泽浅白	色泽青绿	色泽乌黑	根蒂蜷缩	根蒂稍蜷	根蒂硬挺	敲声沉闷	敲声清脆	敲声浊响
0	1	0	1	0	0	0	0	1
0	0	1	1	0	0	1	0	0
0	0	1	1	0	0	0	0	1
0	1	0	1	0	0	1	0	0
1	0	0	1	0	0	0	0	1

纹理模糊	纹理清晰	纹理稍糊	脐部凹陷	脐部平坦	脐部稍凹	触感软粘	触感硬滑	密度
0	1	0	1	0	0	0	1	0.697
0	1	0	1	0	0	0	1	0.774
0	1	0	1	0	0	0	1	0.634
0	1	0	1	0	0	0	1	0.608
0	1	0	1	0	0	0	1	0.556

含糖率	好瓜否	好瓜是
0.460	0	1
0.376	0	1
0.264	0	1
0.318	0	1
0.215	0	1

1.2 神经网络模型

1.2.1 批量梯度下降




```
## [1] "The Train accuracy is :  100 %"
```

```
## [1] "The confusion matrix is : "
```

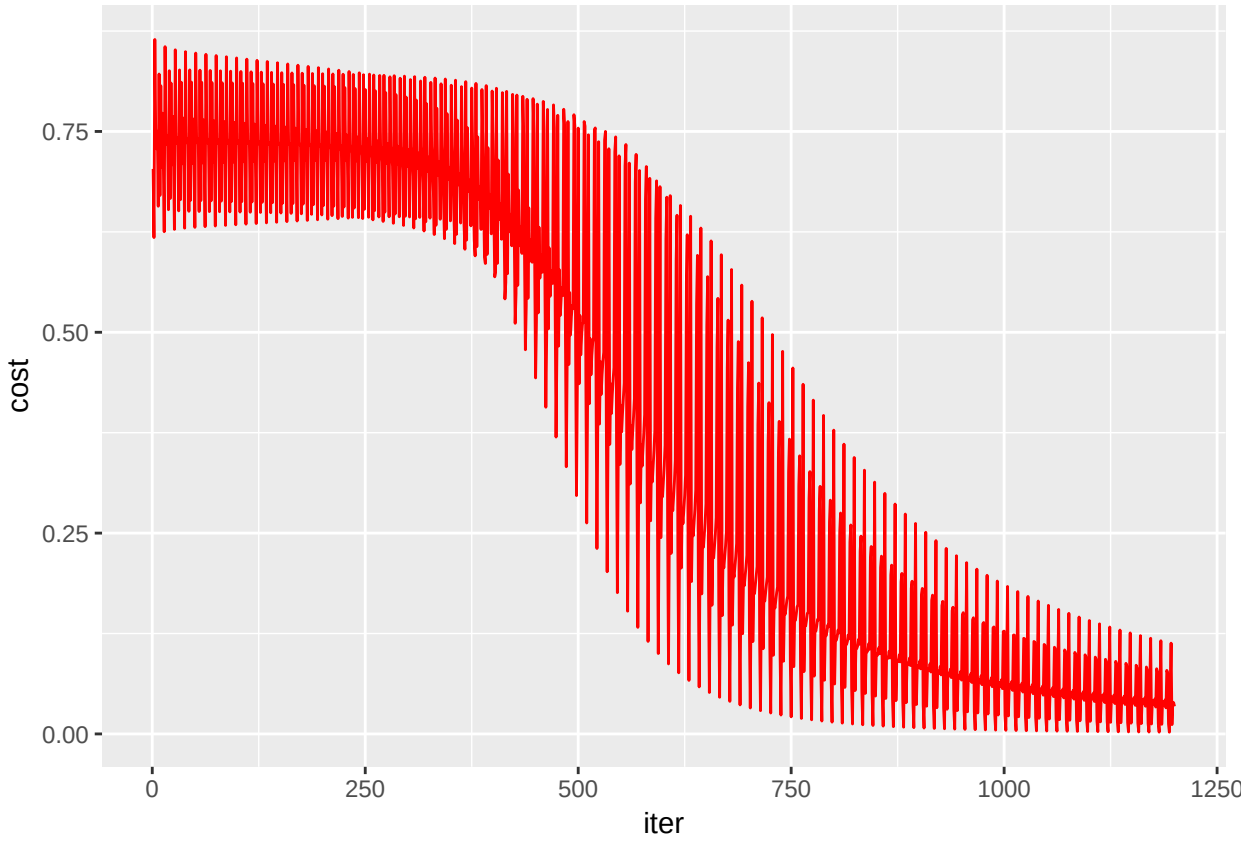
```
##     NewY
## PreY 0 1
##    0 6 0
##    1 0 6
```

```
## [1] "The Test accuracy is :  80 %"
```

```
## [1] "The confusion matrix is : "
```

```
##     NewY
## PreY 0 1
##    0 2 0
##    1 1 2
```

1.2.2 随机梯度下降




```
## [1] "The Train accuracy is :  100 %"

## [1] "The confusion matrix is : "

##      NewY
## PreY 0 1
##    0 6 0
##    1 0 6

## [1] "The Test accuracy is :  80 %"

## [1] "The confusion matrix is : "

##      NewY
## PreY 0 1
##    0 2 0
##    1 1 2
```