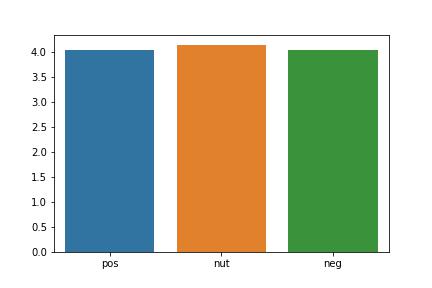
基本数据描述

数据词性

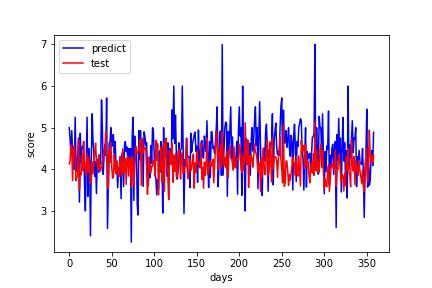
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **attributes** | | | | | |
|  | | 频率 | 百分比 | 有效百分比 | 累积百分比 |
| 有效 | CC | 2 | .6 | .6 | .6 |
| DT | 1 | .3 | .3 | .8 |
| FW | 1 | .3 | .3 | 1.1 |
| IN | 3 | .8 | .8 | 1.9 |
| JJ | 86 | 24.0 | 24.0 | 25.9 |
| JJR | 1 | .3 | .3 | 26.2 |
| JJS | 1 | .3 | .3 | 26.5 |
| MD | 2 | .6 | .6 | 27.0 |
| NN | 205 | 57.1 | 57.1 | 84.1 |
| NNS | 5 | 1.4 | 1.4 | 85.5 |
| PRP$ | 1 | .3 | .3 | 85.8 |
| RB | 12 | 3.3 | 3.3 | 89.1 |
| RBR | 1 | .3 | .3 | 89.4 |
| VB | 6 | 1.7 | 1.7 | 91.1 |
| VBD | 10 | 2.8 | 2.8 | 93.9 |
| VBG | 3 | .8 | .8 | 94.7 |
| VBN | 4 | 1.1 | 1.1 | 95.8 |
| VBP | 12 | 3.3 | 3.3 | 99.2 |
| VBZ | 3 | .8 | .8 | 100.0 |
| 总计 | 359 | 100.0 | 100.0 |  |

各属性与均值的关系

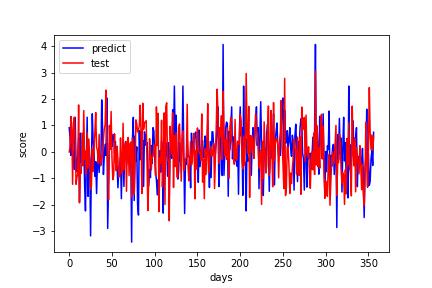


预测

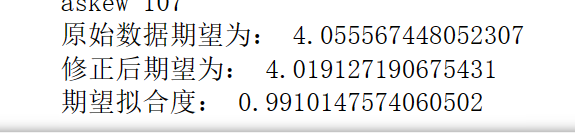
修改前



修改后



测试修改的准确性：取前250个数据求期望 后109个数据测试，期望误差<0.01



测试集和预测集的分布差异

