

Naive Bayes classifier

ساختار و پوشه بندی کل پروژه :

پوشه src : شامل تمامی کد های اجرایی است.
پوشه data : شامل تمامی داده های جمع آوری شده یا ساخته شده است .
پوشه tools : شامل ابزار های استفاده شده در کد برنامه است.
پوشه analyze : شامل نمودار ها و نتایج بدست آمده از در هر فاز و نیز داکيومنت ها است.

فایل های مربوط به این فاز :

p2_naive_bayes.py

مکانیزم جدا سازی داده ی تست و ترین :

با استفاده از دستور زیر داده ی تست را به صورت رندوم از لیست کامن ها جدا میکنیم و در لیست تست وارد میکنیم.

```
tc_celebrity_cms = []
tc_friendly_cms = []
test_case_count = 200
for _ in range(test_case_count):
    tc_celebrity_cms.append(celebrity_cms.pop(random.randint(0,
len(celebrity_cms) - 1)))
    tc_friendly_cms.append(friendly_cms.pop(random.randint(0,
len(friendly_cms) - 1)))
```

و مشابه زیر خروجی را برای داده های تست بدست می آوریم

```
print ("without smothing")
print("celebrity: precision is : ", precision(tc_celebrity_cms,
tc_friendly_cms, "celebrity"), "recall is:",
recall(tc_celebrity_cms, "celebrity"))
```

```
print("friendly: precision is : ", precision(tc_friendly_cms,  
tc_celebrity_cms, "friendly"), "recall is:",  
recall(tc_friendly_cms, "friendly"))
```

نحوه اجرای کد :

Python3 p2_naive_bayes.py

نتایج خروجی برای یک تست کیس رندوم با تست ریت ۱۰ کل ورودی :

without smoothing

celebrity: precision is : 0.9428571428571428 recall is: 0.165

friendly: precision is : 0.5424657534246575 recall is: 0.99

with smothing

celebrity: precision is : 0.8018018018018018 recall is: 0.445

friendly: precision is : 0.615916955017301 recall is: 0.89