مبانی هوش محاسباتی گزارش تمرین اول

شرح كلى:

برای clustering این دیتاست به ۳ روش kmeans, dbscan و kmeans این دیتاست به ۳ روش clustering این دیتاست. العام است. العام است. العام العام

نحوه محاسبه rand index:

این میزان با استفاده از فرمول زیر محاسبه می شود.

rand index =
$$\frac{TP+TN}{TP+FP+TN+FN}$$

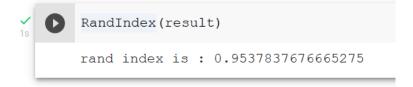
Rand Index

```
def RandIndex(clusteredData):
    TP_and_FP = calculateTP_FP(clusteredData)
    TP = calculateTP(clusteredData)
    FN = 1800 - TP
    totalClusteredData = combination(410 - np.count_nonzero(clusteredData == -1))
    TN = totalClusteredData - (FN + TP_and_FP)
    randIndex = (TP + TN) / totalClusteredData
    print(f"rand index is : {randIndex}")
```

:rand index نتایج معیار

Kmeans •

Algorithm Evaluation



Dbscan •

Algorithm Evaluation

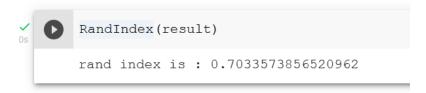
```
RandIndex(result)

rand index is: 0.3277857484457197
```

Agglomerative •

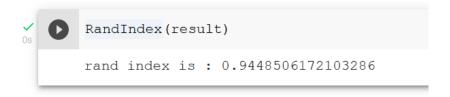
Single Link o

Algorithm Evaluation



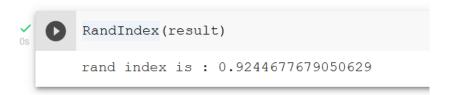
Complete Link o

Algorithm Evaluation



Average Link o

Algorithm Evaluation



همانطور که مشاهده می کنید الگوریتم dbscan امتیاز بسیار پایینی در این معیار دریافت کرده زیرا بسیاری از داده ها را به شکل نویز تشخیص داده است.

در ضمن، نتایج کلی کلاسترینگ پس از اتمام الگوریتم در پوشه هایی با نام هر کلاستر به شکل زیر ذخیره می شود. Paths شامل آدرس تمامی عکس هاست. و result نتایج کلاسترینگ را شامل می شود.

Collecting Images

محیا احسانی مهر ۹۸۱۲۷۶۲۳۲۷