UAS KONSEP APLIKASI DATA MAINING ASSOCIATION RULE (MENGHITUNG SUPPORT DAN CONFIDANCE) MENGGUNAKAN JUPYTER NETBOOK

Nama : Himatus Yulvi A.S

Jurusan :Sistem Infomasi

Nim :17.51.0005

KEMENTRIAN RISET DAN TEKNOLOGI PENDIDIKAN TINGGI
SEKOLAH TINGGI MANAJEMEN INFORMATIKA KOMPUTER
PRADNYA PARAMITA
MALANG
2020

Dalam aturan asosiasi terdapat 3 macam pengukuran penting, yaitu Support, Confidence, dan Lift Rasio. 3 asosioasi di atas sering di gunakan di **Association Rule.**

- Support adalah pencarian jumlah tranksaksi yang mengandung item berbanding dengan total tranksaksi Dan indikasi sering muncul di dataset.

- Confidence adalah nilai kepastian atau kuatnya hubungan antar item dalam aturan asosiasi.(indikasi seberapa sering aturan itu terbukti benar)

- Nilai lift rasio merupakan suatu ukuran dalam mengetahui kekuatan suatu aturan asosiasi. (indikasi nilai pengamatan Support , diharapkan jika dua aturan itu independen)

SOAL:

3. Lakukan clasifikasi dengan menggunakan Algoritam Association Rules dalam soal No. 3 dan hitunglah apabila:

a. min_support=0.2, min_confidence=0.2, min_lift=0.2, min_lenght=2

b. Berapa association rules yang dapat terbentuk dari soal 3a. tersebut?

c. Lakukan export data ke dalam excel terhadap hasil tersebut dan lakukan analisanya!

Output:

```
In [2]: store_data= pd.read_excel('C:/Users/Yulvavi/Documents/KULIAH/Himatus Yulvi/Sistem Data maining/dataset_soal No. 3.xls')
```

```
In [3]: store_data.head()
```

Out[3]:

	Item1	Item2	Item3	Item4	Item5	Item6	Item7	Item8	Item9	Item10
0	burgers	meatballs	eggs	low fat yogurt	NaN	mineral water	salmon	low fat yogurt	NaN	mineral water
1	chutney	low fat yogurt	NaN	whole wheat pasta	french fries	mineral water	salmon	whole wheat pasta	french fries	mineral water
2	turkey	whole wheat pasta	french fries	soup	light cream	shallot	NaN	soup	light cream	shallot
3	mineral water	soup	light cream	frozen vegetables	spaghetti	green tea	NaN	frozen vegetables	spaghetti	green tea
4	low fat yogurt	frozen vegetables	spaghetti	french fries	eggs	chocolate	frozen smoothie	french fries	eggs	chocolate

```
In [4]: store_data.tail()
```

Out[4]:

	Item1	Item2	Item3	Item4	Item5	Item6	Item7	Item8	Item9	Item10
2049	burgers	eggs	french fries	fresh tuna	spaghetti	olive oil	clothes accessories	turkey	eggs	french fries
2050	burgers	eggs	frozen smoothie	french wine	eggs	french fries	energy drink	french fries	NaN	chocolate
2051	whole wheat pasta	cake	melons	champagne	pancakes	light mayo	soup	chocolate	milk	herb & pepper
2052	ground beef	tomato sauce	spaghetti	red wine	honey	hot dogs	turkey	herb & pepper	whole wheat pasta	mineral water
2053	burgers	eggs	frozen smoothie	milk	bacon	eggs	french fries	mineral water	avocado	cookies

a. min_support=0.2, min_confidence=0.2, min_lift=0.2, min_lenght=2

```
In [6]: records = []
        for i in range (0, 2054):
            records.append([str(store_data.values[i,j])for j in range (0,10)])
```

```
In [7]: association_rules = apriori(records, min_support=0.2, min_confidence=0.2, min_lift=0.2, min_length=2)
        association_results = list(association_rules)
```

b. Berapa association rules yang dapat terbentuk dari soal 3a. tersebut?

```
In [8]: print(len(association_results))
```

```
61
```

```
In [9]: print(association_results[0])
```

```
RelationRecord(items=frozenset({'avocado'}), support=0.314508276533593, ordered_statistics=[OrderedStatistic(items_base=frozens
et(), items_add=frozenset({'avocado'}), confidence=0.314508276533593, lift=1.0)])
```

```
In [14]: print (items)
```

```
[frozenset({'eggs', 'french fries', 'turkey', 'clothes accessories'}), 0.24196689386562803, [OrderedStatistic(items_base=frozen
set(), items_add=frozenset({'eggs', 'french fries', 'turkey', 'clothes accessories'}), confidence=0.24196689386562803, lift=1.
0), OrderedStatistic(items_base=frozenset({'clothes accessories'}), items_add=frozenset({'eggs', 'french fries', 'turkey'}), co
nfidence=0.7120343839541546, lift=2.683520412186851), OrderedStatistic(items_base=frozenset({'eggs'}), items_add=frozenset({'fr
ench fries', 'turkey', 'clothes accessories'}), confidence=0.5902612826603325, lift=1.8709825225066714), OrderedStatistic(items
_base=frozenset({'french fries'}), items_add=frozenset({'eggs', 'turkey', 'clothes accessories'}), confidence=0.369516728624535
3, lift=1.4795075255259171), OrderedStatistic(items_base=frozenset({'turkey'}), items_add=frozenset({'eggs', 'french fries', 'c
lothes accessories'}), confidence=0.458910433979686, lift=1.8889820268422346), OrderedStatistic(items_base=frozenset({'eggs',
'clothes accessories'}), items_add=frozenset({'french fries', 'turkey'}), confidence=0.9631782945736435, lift=2.59628374941504
4), OrderedStatistic(items_base=frozenset({'french fries', 'clothes accessories'}), items_add=frozenset({'eggs', 'turkey'}), co
nfidence=0.7473684210526316, lift=2.6240935672514616), OrderedStatistic(items_base=frozenset({'turkey', 'clothes accessorie
s'}), items_add=frozenset({'eggs', 'french fries'}), confidence=0.7373887240356083, lift=2.0384878050728656), OrderedStatistic
(items_base=frozenset({'eggs', 'french fries'}), items_add=frozenset({'turkey', 'clothes accessories'}), confidence=0.668909825
0336473, lift=2.0384878050728656), OrderedStatistic(items_base=frozenset({'eggs', 'turkey'}), items_add=frozenset({'french frie
s', 'clothes accessories'}), confidence=0.8495726495726494, lift=2.6240935672514616), OrderedStatistic(items_base=frozenset({'f
rench fries', 'turkey'}), items_add=frozenset({'eggs', 'clothes accessories'}), confidence=0.6522309711286088, lift=2.596283749
415044), OrderedStatistic(items_base=frozenset({'eggs', 'french fries', 'clothes accessories'}), items_add=frozenset({'turke
y'}), confidence=0.9959919839679358, lift=1.8889820268422346), OrderedStatistic(items_base=frozenset({'eggs', 'turkey', 'clothe
s accessories'}), items_add=frozenset({'french fries'}), confidence=0.9688109161793371, lift=1.4795075255259171), OrderedStatis
tic(items_base=frozenset({'french fries', 'turkey', 'clothes accessories'}), items_add=frozenset({'eggs'}), confidence=0.766975
3086419753, lift=1.8709825225066714), OrderedStatistic(items_base=frozenset({'eggs', 'french fries', 'turkey'}), items_add=froz
enset({'clothes accessories'}), confidence=0.9119266055045872, lift=2.683520412186851)]]
```

Script untuk menampilkan support, confidence, dan lift, maka jalankan script dibawah ini:

```
In [11]: results = []
for item in association_results:
    pair = item
    items = [x for x in pair]

    value0 = str(items[0])
    value1 = str(items[1])
    value2 = str(item[1])[:10]
    value3 = str(item[2][0][2])[:10]
    value4 = str(item[2][0][3])[:10]

    rows = (value0, value1, value2, value3, value4)

    results.append(rows)

    label = ['title1'], ['title2'],['support'], ['confidence'],['lift']

    store_suggestion = pd.DataFrame.from_records(results,columns=label)

    print(store_suggestion)
```

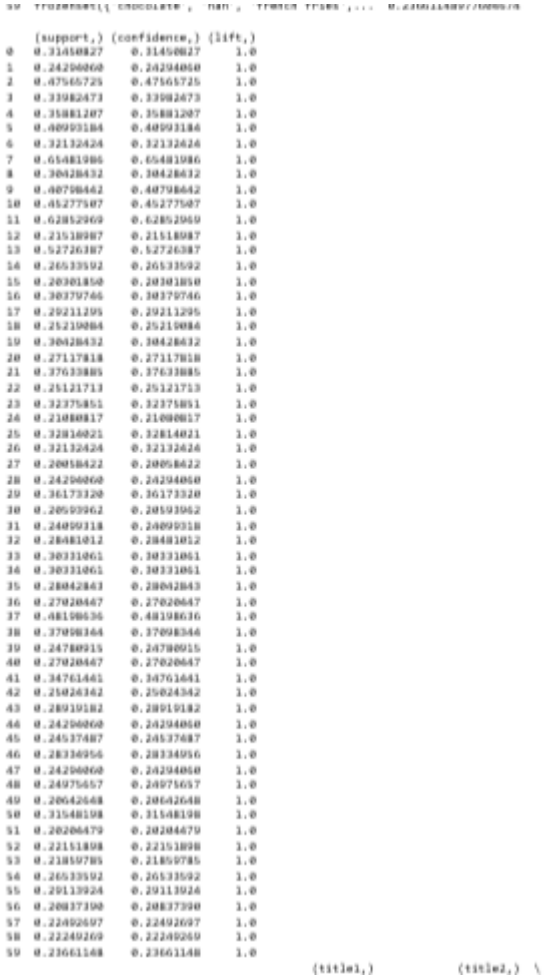
```
                (title1,)          (title2,)  (support,) (confidence,) (lift,)
0  frozenset({'avocado'})  0.314508276533593  0.31450827    0.31450827     1.0
                (title1,)           (title2,)  (support,) (confidence,)  \
0  frozenset({'avocado'})    0.314508276533593  0.31450827    0.31450827
1  frozenset({'burgers'})  0.24294060370009737  0.24294060    0.24294060

  (lift,)
0     1.0
1     1.0
                  (title1,)           (title2,)  (support,) (confidence,)  \
0    frozenset({'avocado'})    0.314508276533593  0.31450827    0.31450827
1    frozenset({'burgers'})  0.24294060370009737  0.24294060    0.24294060
2  frozenset({'chocolate'})   0.4756572541382668  0.47565725    0.47565725
```

	(confidence,)	(lift,)	(title1,)	(title2,)	(support,)	\
0	0.31450827	1.0	frozenset({'avocado'})	0.314508276533593	0.31450827	
1	0.24294060	1.0	frozenset({'burgers'})	0.24294060370009737	0.24294060	
2	0.47565725	1.0	frozenset({'chocolate'})	0.4756572541382668	0.47565725	
3	0.33982473	1.0	frozenset({'clothes accessories'})	0.33982473222979553	0.33982473	
4	0.35881207	1.0	frozenset({'cookies'})	0.3588120740019474	0.35881207	
5	0.40993184	1.0	frozenset({'eggs'})	0.40993184031158714	0.40993184	
6			frozenset({'energy drink'})	0.3213242453748783	0.32132424	

	(confidence,)	(lift,)	(title1,)	(title2,)	(support,)	\
0	0.31450827	1.0	frozenset({'avocado'})	0.314508276533593	0.31450827	
1	0.24294060	1.0	frozenset({'burgers'})	0.24294060370009737	0.24294060	
2	0.47565725	1.0	frozenset({'chocolate'})	0.4756572541382668	0.47565725	
3	0.33982473	1.0	frozenset({'clothes accessories'})	0.33982473222979553	0.33982473	
4	0.35881207	1.0	frozenset({'cookies'})	0.3588120740019474	0.35881207	
5	0.40993184	1.0	frozenset({'eggs'})	0.40993184031158714	0.40993184	
6	0.32132424	1.0	frozenset({'energy drink'})	0.3213242453748783	0.32132424	
7			frozenset({'french fries'})	0.6548198636806232	0.65481986	

1	0.24294060	0.24294060	1.0	
2	0.47565725	0.47565725	1.0	
3	0.33982473	0.33982473	1.0	
4	0.35881207	0.35881207	1.0	
5	0.40993184	0.40993184	1.0	
6	0.32132424	0.32132424	1.0	
7	0.65481986	0.65481986	1.0	
8	0.30428432	0.30428432	1.0	
9	0.40798442	0.40798442	1.0	
10	0.45277507	0.45277507	1.0	
11	0.62852969	0.62852969	1.0	
12	0.21518987	0.21518987	1.0	
13	0.52726387	0.52726387	1.0	
14	0.26533592	0.26533592	1.0	
15	0.20301850	0.20301850	1.0	
16	0.30379746	0.30379746	1.0	
17	0.29211295	0.29211295	1.0	
18	0.25219084	0.25219084	1.0	
19	0.30428432	0.30428432	1.0	
20	0.27117818	0.27117818	1.0	
21	0.37633885	0.37633885	1.0	
22	0.25121713	0.25121713	1.0	
23	0.32375851	0.32375851	1.0	
24	0.21080817	0.21080817	1.0	
25	0.32814021	0.32814021	1.0	
26	0.32132424	0.32132424	1.0	
27	0.20058422	0.20058422	1.0	
28	0.24294060	0.24294060	1.0	
29	0.36173320	0.36173320	1.0	
30	0.20593962	0.20593962	1.0	
31	0.24099318	0.24099318	1.0	
32	0.28481012	0.28481012	1.0	
33	0.30331061	0.30331061	1.0	
34	0.30331061	0.30331061	1.0	
35	0.28042843	0.28042843	1.0	
36	0.27020447	0.27020447	1.0	



```
    (support,) (confidence,) (lift,)
0   0.31458827    0.31458827     1.0
1   0.24294060    0.24294060     1.0
2   0.47565725    0.47565725     1.0
3   0.33982473    0.33982473     1.0
4   0.35881207    0.35881207     1.0
5   0.40993184    0.40993184     1.0
6   0.32132424    0.32132424     1.0
7   0.65481986    0.65481986     1.0
8   0.30428432    0.30428432     1.0
9   0.40798442    0.40798442     1.0
10  0.45277507    0.45277507     1.0
11  0.62852969    0.62852969     1.0
12  0.21518987    0.21518987     1.0
13  0.52726387    0.52726387     1.0
14  0.26533592    0.26533592     1.0
15  0.20301850    0.20301850     1.0
16  0.30379746    0.30379746     1.0
17  0.29211295    0.29211295     1.0
18  0.25219084    0.25219084     1.0
19  0.30428432    0.30428432     1.0
20  0.27117818    0.27117818     1.0
21  0.37633885    0.37633885     1.0
22  0.25121713    0.25121713     1.0
23  0.32375851    0.32375851     1.0
24  0.21080817    0.21080817     1.0
25  0.32814021    0.32814021     1.0
26  0.32132424    0.32132424     1.0
27  0.20058422    0.20058422     1.0
28  0.24294060    0.24294060     1.0
29  0.36173320    0.36173320     1.0
30  0.20593962    0.20593962     1.0
31  0.24099318    0.24099318     1.0
32  0.28481012    0.28481012     1.0
33  0.30331061    0.30331061     1.0
34  0.30331061    0.30331061     1.0
35  0.28042843    0.28042843     1.0
36  0.27020447    0.27020447     1.0
37  0.48198636    0.48198636     1.0
38  0.37098344    0.37098344     1.0
39  0.24780915    0.24780915     1.0
40  0.27020447    0.27020447     1.0
41  0.34761441    0.34761441     1.0
42  0.25024342    0.25024342     1.0
43  0.28919182    0.28919182     1.0
44  0.24294060    0.24294060     1.0
45  0.24537487    0.24537487     1.0
46  0.28334956    0.28334956     1.0
47  0.24294060    0.24294060     1.0
48  0.24975657    0.24975657     1.0
49  0.20642648    0.20642648     1.0
50  0.31548198    0.31548198     1.0
51  0.20204479    0.20204479     1.0
52  0.22151898    0.22151898     1.0
53  0.21859785    0.21859785     1.0
54  0.26533592    0.26533592     1.0
55  0.29113924    0.29113924     1.0
56  0.20837390    0.20837390     1.0
57  0.22492697    0.22492697     1.0
58  0.22249269    0.22249269     1.0
59  0.23661148    0.23661148     1.0
                                          (title1,)          (title2,) \
```

c. Lakukan export data ke dalam excel terhadap hasil tersebut dan lakukan analisanya!

Selanjutnya jalankan script berikut ini untuk menampilkan kesimpulan

```
store_suggestion.describe()
```

Output:

Out[12]:

	(title1,)	(title2,)	(support,)	(confidence,)	(lift,)
count	61	61	61	61	61
unique	61	53	53	53	1
top	frozenset({'chocolate', 'nan'})	0.24294060370009737	0.24294060	0.24294060	1.0
freq	1	4	4	4	61

Selanjutnya meneksport hasilnya ke bentuk excel dengan menggunakan script berikut ini :

Kesimpulan:

Algoritma Association rule mining seperti Apriori sangat fleksibel dan mudah di gunakan karena sangat sederhana . Mereka mudah diimplementasikan dan memiliki kemampuan menjelaskan yang tinggi. Pada tugas ini saya mencoba (min_support=0.2, min_confidence=0.2, min_lift=0.2, min_lenght=2) dan mendapatkan 60 list dan yang tertinggi adalah frozenset({'chocolate', 'nan'}) dengan Support 0.24294060dan Confidence 0.24294060