Analitika Data I

Association Rule (Apriori Algorithm)

Putra Pandu Adikara

Fakultas Ilmu Komputer
Universitas Brawijaya

Alat Bantu 1: mlextend

- Menggunakan library mlxtend
  - Instalasi:
    - Anaconda: `conda install mlxtend`
    - Vanilla Python dgn PIP: `pip install mlextend`

- Pemanggilan:
  - `from mlxtend.frequent_patterns import apriori`
  - `from mlxtend.frequent_patterns import fpmax, fpgrowth #bisa juga menggunakan`
  - `from mlxtend.frequent_patterns import association_rules`

Pembuatan Association Rule dari Frequent Itemsets

```python
import pandas as pd
from mlxtend.preprocessing import TransactionEncoder
from mlxtend.frequent_patterns import apriori, fpmax, fpgrowth
```

```python
dataset = [['Milk', 'Onion', 'Nutmeg', 'Kidney Beans', 'Eggs', 'Yogurt'],
           ['Dill', 'Onion', 'Nutmeg', 'Kidney Beans', 'Eggs', 'Yogurt'],
           ['Milk', 'Apple', 'Kidney Beans', 'Eggs'],
           ['Milk', 'Unicorn', 'Corn', 'Kidney Beans', 'Yogurt'],
           ['Corn', 'Onion', 'Onion', 'Kidney Beans', 'Ice cream', 'Eggs']]

te = TransactionEncoder()
te_ary = te.fit(dataset).transform(dataset)
df = pd.DataFrame(te_ary, columns=te.columns_)

frequent_itemsets = fpgrowth(df, min_support=0.6, use_colnames=True)
### alternatively:
#frequent_itemsets = apriori(df, min_support=0.6, use_colnames=True)
#frequent_itemsets = fpmax(df, min_support=0.6, use_colnames=True)
```

FILKOM UB

```
1 frequent_itemsets
```

	support	itemsets
0	1.0	(Kidney Beans)
1	0.8	(Eggs)
2	0.6	(Yogurt)
3	0.6	(Onion)
4	0.6	(Milk)
5	0.8	(Kidney Beans, Eggs)
6	0.6	(Yogurt, Kidney Beans)
7	0.6	(Onion, Eggs)
8	0.6	(Kidney Beans, Onion)
9	0.6	(Kidney Beans, Onion, Eggs)
10	0.6	(Milk, Kidney Beans)

Pembuatan Rule dan Kriteria Seleksi

- Seleksi dengan metric confidence

```
1 from mlxtend.frequent_patterns import association_rules
2
3 association_rules(frequent_itemsets, metric="confidence", min_threshold=0.7)
```

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(Kidney Beans)	(Eggs)	1.0	0.8	0.8	0.80	1.00	0.00	1.0
1	(Eggs)	(Kidney Beans)	0.8	1.0	0.8	1.00	1.00	0.00	inf
2	(Yogurt)	(Kidney Beans)	0.6	1.0	0.6	1.00	1.00	0.00	inf
3	(Onion)	(Eggs)	0.6	0.8	0.6	1.00	1.25	0.12	inf
4	(Eggs)	(Onion)	0.8	0.6	0.6	0.75	1.25	0.12	1.6
5	(Onion)	(Kidney Beans)	0.6	1.0	0.6	1.00	1.00	0.00	inf
6	(Onion, Kidney Beans)	(Eggs)	0.6	0.8	0.6	1.00	1.25	0.12	inf
7	(Kidney Beans, Eggs)	(Onion)	0.8	0.6	0.6	0.75	1.25	0.12	1.6
8	(Onion, Eggs)	(Kidney Beans)	0.6	1.0	0.6	1.00	1.00	0.00	inf

Pembuatan Rule dan Kriteria Seleksi

- Seleksi dengan metric lift

```
1 rules = association_rules(frequent_itemsets, metric="lift", min_threshold=1.2)
2 rules
```

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(Onion)	(Eggs)	0.6	0.8	0.6	1.00	1.25	0.12	inf
1	(Eggs)	(Onion)	0.8	0.6	0.6	0.75	1.25	0.12	1.6
2	(Onion, Kidney Beans)	(Eggs)	0.6	0.8	0.6	1.00	1.25	0.12	inf
3	(Kidney Beans, Eggs)	(Onion)	0.8	0.6	0.6	0.75	1.25	0.12	1.6
4	(Onion)	(Kidney Beans, Eggs)	0.6	0.8	0.6	1.00	1.25	0.12	inf
5	(Eggs)	(Onion, Kidney Beans)	0.8	0.6	0.6	0.75	1.25	0.12	1.6

- Membuat Antecedent dan Consequent

```
1 rules["antecedent_len"] = rules["antecedents"].apply(lambda x: len(x))
2 rules
```

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	antecedent_len
0	(Onion)	(Eggs)	0.6	0.8	0.6	1.00	1.25	0.12	inf	1
1	(Eggs)	(Onion)	0.8	0.6	0.6	0.75	1.25	0.12	1.6	1
2	(Onion, Kidney Beans)	(Eggs)	0.6	0.8	0.6	1.00	1.25	0.12	inf	2
3	(Kidney Beans, Eggs)	(Onion)	0.8	0.6	0.6	0.75	1.25	0.12	1.6	2
4	(Onion)	(Kidney Beans, Eggs)	0.6	0.8	0.6	1.00	1.25	0.12	inf	1
5	(Eggs)	(Onion, Kidney Beans)	0.8	0.6	0.6	0.75	1.25	0.12	1.6	1

Pembuatan Rule dan Kriteria Seleksi

- Memilih rule sesuai kriteria dan panjang antecedent

```
rules[ (rules['antecedent_len'] >= 2) &
       (rules['confidence'] > 0.75) &
       (rules['lift'] > 1.2) ]
```

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	antecedent_len
2	(Onion, Kidney Beans)	(Eggs)	0.6	0.8	0.6	1.0	1.25	0.12	inf	2

Pembuatan Rule dan Kriteria Seleksi

- Memilih antecedent tertentu (urutan tidak memengaruhi karena set)

```
1 rules[rules['antecedents'] == {'Eggs', 'Kidney Beans'}]
2
```

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	antecedent_len
3	(Kidney Beans, Eggs)	(Onion)	0.8	0.6	0.6	0.75	1.25	0.12	1.6	2

```
1
```

Alat Bantu 1: apyori

- Cek: https://pypi.org/project/apyori/1.0.0/