pt-3

November 5, 2020

0.1 Practice №3

[1]:
```
import pandas as pd
from mlxtend.preprocessing import TransactionEncoder
from mlxtend.frequent_patterns import apriori, fpgrowth
```

[2]:
```
dataset = [list('ABCD'), list('ACDF'), list('ACDEG'), list('ABDF'),␣
↪list('BCG'), list('DFG'), list('ABG'), list('CDFG')]
```

[3]:
```
te = TransactionEncoder()
te_ary = te.fit(dataset).transform(dataset)
```

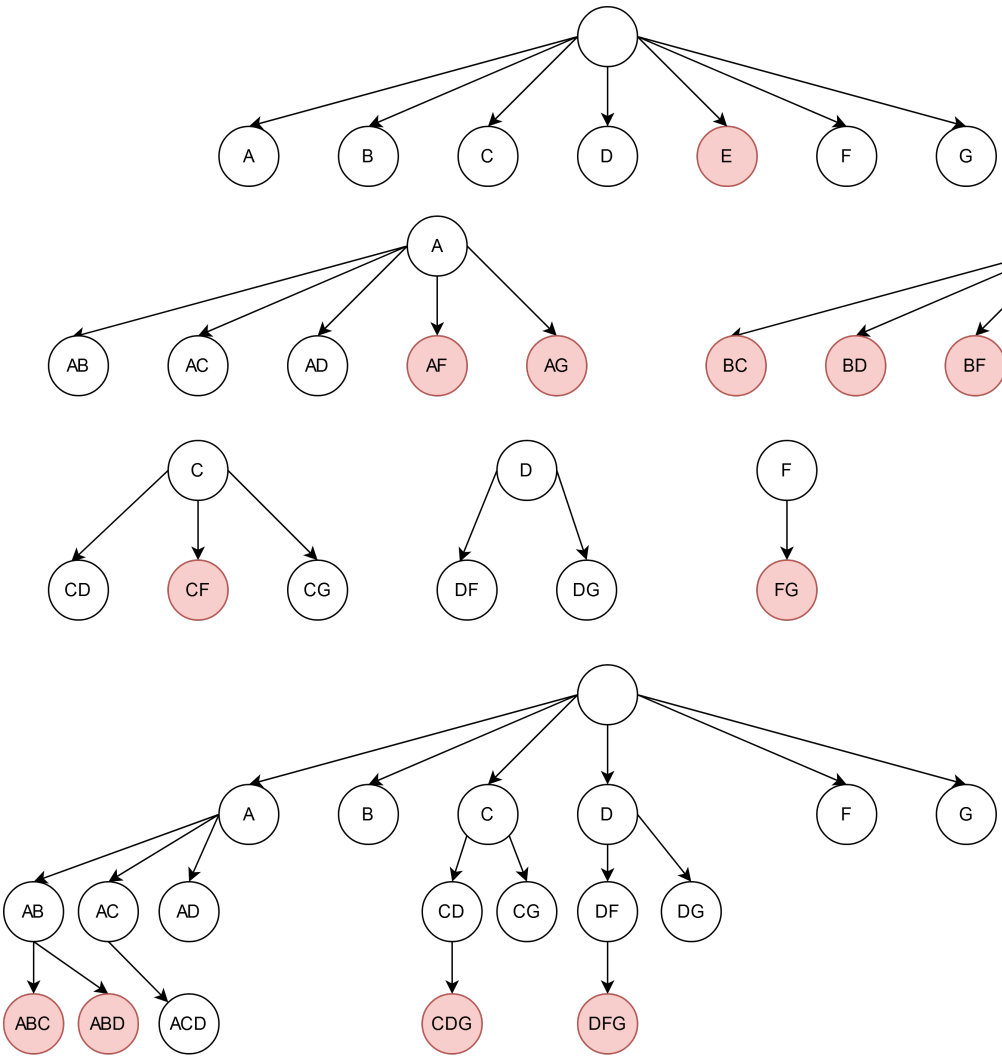
[4]:
```
df = pd.DataFrame(te_ary, columns=['A', 'B', 'C', 'D', 'E', 'F', 'G'], index=[i␣
↪for i in range(1, 9)])
df
```

[4]:

	A	B	C	D	E	F	G
1	True	True	True	True	False	False	False
2	True	False	True	True	False	True	False
3	True	False	True	True	True	False	True
4	True	True	False	True	False	True	False
5	False	True	True	False	False	False	True
6	False	False	False	True	False	True	True
7	True	True	False	False	False	False	True
8	False	False	True	True	False	True	True

0.1.1 Task №1

Apriori



[5]:
```
results = apriori(df, min_support=3/8, use_colnames=True)
results['length'] = results['itemsets'].apply(lambda x: len(x))
results
```

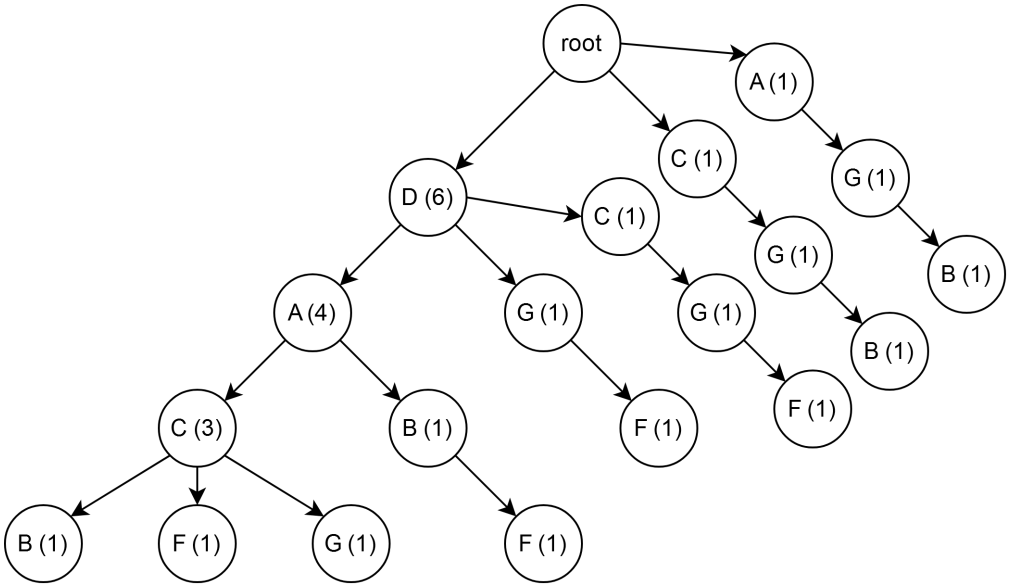
[5]:

	support	itemsets	length
0	0.625	(A)	1
1	0.500	(B)	1
2	0.625	(C)	1
3	0.750	(D)	1
4	0.500	(F)	1
5	0.625	(G)	1
6	0.375	(A, B)	2
7	0.375	(A, C)	2
8	0.500	(A, D)	2
9	0.500	(D, C)	2
10	0.375	(G, C)	2
11	0.500	(D, F)	2

```
12    0.375     (G, D)       2
13    0.375  (A, D, C)       3
```

FPGrowth

id	itemset
1	DACB
2	DACF
3	DACG
4	DABF
5	CGB
6	DGF
7	AGB
8	DCGF



[6]:
```
results = fpgrowth(df, min_support=2/8, use_colnames=True)
results['length'] = results['itemsets'].apply(lambda x: len(x))
results
```

[6]:
```
    support   itemsets  length
0     0.750        (D)       1
1     0.625        (C)       1
2     0.625        (A)       1
3     0.500        (B)       1
4     0.500        (F)       1
5     0.625        (G)       1
6     0.500     (D, C)       2
7     0.375     (G, C)       2
```

```
8     0.250  (G, D, C)       3
9     0.500     (A, D)       2
10    0.375     (A, C)       2
11    0.250     (A, G)       2
12    0.375  (A, D, C)       3
13    0.375     (A, B)       2
14    0.250     (C, B)       2
15    0.250     (D, B)       2
16    0.250     (G, B)       2
17    0.250  (A, D, B)       3
18    0.500     (D, F)       2
19    0.250     (A, F)       2
20    0.250     (F, C)       2
21    0.250     (G, F)       2
22    0.250  (A, D, F)       3
23    0.250  (D, C, F)       3
24    0.250  (G, D, F)       3
25    0.375     (G, D)       2
```

0.1.2 Task №2

{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11}; Size = 2^11 - 1 = 2047

High order itemset

id	itemset
1	12 14 15
2	12 14 13 15
3	12 14 13 15
4	14 15
5	12 14 13 15
6	12 14 13 15
7	12 14 13 15
8	12 14 13 15

Support 7/8

itemset	support
12	7/8
14	1
15	1
12 14	7/8
12 15	7/8
14 15	1

[ ]: