Apriori Association Rule Learning Working Copy

July 15, 2024

1 Apriori

1.1 Importing libraries

[2]:
```
!pip install apyori
```

```
Collecting apyori
  Downloading https://files.pythonhosted.org/packages/5e/62/5ffde5c473ea4b033490
617ec5caa80d59804875ad3c3c57c0976533a21a/apyori-1.1.2.tar.gz
Building wheels for collected packages: apyori
  Building wheel for apyori (setup.py) … done
  Created wheel for apyori: filename=apyori-1.1.2-cp36-none-any.whl size=5977
sha256=5798642d22453aa96e676f40db040f645b265f526419c1a786bcd3cab9e88aa5
  Stored in directory: /root/.cache/pip/wheels/5d/92/bb/474bbadbc8c0062b9eb168f6
9982a0443263f8ab1711a8cad0
Successfully built apyori
Installing collected packages: apyori
Successfully installed apyori-1.1.2
```

[3]:
```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

1.2 Importing dataset

[4]:
```
dataset = pd.read_csv('Big Basket.com Cart.csv', header = None)
transactions = []

for i in range(0, 7219):
  transactions.append([str(dataset.values[i,j]) for j in range(0, 20)])
```

1.3 Apriori Training on Dataset

[5]:
```
from apyori import apriori
rules = apriori(transactions = transactions, min_support = 0.003,␣
 ↪min_confidence = 0.2, min_lift = 3, min_length = 2, max_length = 2)
```

1.4 Visualizing

1.4.1 Raw Results

[6]:
```
results = list(rules)
```

[7]:
```
results
```

[7]:
```
[RelationRecord(items=frozenset({'burgers', 'almonds'}),
support=0.005402410306136584,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'almonds'}),
items_add=frozenset({'burgers'}), confidence=0.26530612244897955,
lift=3.0594966421073218)]),
 RelationRecord(items=frozenset({'buns', 'paneer'}),
support=0.016068707577226764,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'buns'}),
items_add=frozenset({'paneer'}), confidence=0.32402234636871513,
lift=3.2852771326344867)]),
 RelationRecord(items=frozenset({'chicken', 'ginger garlic paste'}),
support=0.004571270259038648,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'ginger garlic
paste'}), items_add=frozenset({'chicken'}), confidence=0.29203539823008845,
lift=4.791371681415929)]),
 RelationRecord(items=frozenset({'fruit basket', 'vegetable basket'}),
support=0.0055409336473195734,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'fruit basket'}),
items_add=frozenset({'vegetable basket'}), confidence=0.29411764705882354,
lift=3.081618714249125)]),
 RelationRecord(items=frozenset({'ginger garlic paste', 'olive oil'}),
support=0.003324560188391744,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'ginger garlic
paste'}), items_add=frozenset({'olive oil'}), confidence=0.21238938053097342,
lift=3.2346813039094875)]),
 RelationRecord(items=frozenset({'kissan puree', 'paneer'}),
support=0.005402410306136584,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'kissan puree'}),
items_add=frozenset({'paneer'}), confidence=0.3786407766990291,
lift=3.8390558525144542)]),
 RelationRecord(items=frozenset({'pasta', 'maggi'}),
support=0.0034630835295747335,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'pasta'}),
items_add=frozenset({'maggi'}), confidence=0.24752475247524755,
lift=3.4562498803071797)]),
 RelationRecord(items=frozenset({'pasta', 'mushroom'}),
support=0.005402410306136584,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'pasta'}),
items_add=frozenset({'mushroom'}), confidence=0.38613861386138615,
lift=4.873312331233124)]),
```

```
 RelationRecord(items=frozenset({'mushroom', 'pepper spray'}),
support=0.005817980329685552,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'pepper spray'}),
items_add=frozenset({'mushroom'}), confidence=0.3021582733812949,
lift=3.8134275796146295)]),
 RelationRecord(items=frozenset({'neckrest', 'trolly bag'}),
support=0.0034630835295747335,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'neckrest'}),
items_add=frozenset({'trolly bag'}), confidence=0.2717391304347826,
lift=5.620873302603712)]),
 RelationRecord(items=frozenset({'whole wheat pasta', 'olive oil'}),
support=0.007757307106247403,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'whole wheat
pasta'}), items_add=frozenset({'olive oil'}), confidence=0.26666666666666666,
lift=4.061322081575246)])]
```

[8]:
```
print(results)
```

```
[RelationRecord(items=frozenset({'burgers', 'almonds'}),
support=0.005402410306136584,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'almonds'}),
items_add=frozenset({'burgers'}), confidence=0.26530612244897955,
lift=3.0594966421073218)]), RelationRecord(items=frozenset({'buns', 'paneer'}),
support=0.016068707577226764,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'buns'}),
items_add=frozenset({'paneer'}), confidence=0.32402234636871513,
lift=3.2852771326344867)]), RelationRecord(items=frozenset({'chicken', 'ginger
garlic paste'}), support=0.004571270259038648,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'ginger garlic
paste'}), items_add=frozenset({'chicken'}), confidence=0.29203539823008845,
lift=4.791371681415929)]), RelationRecord(items=frozenset({'fruit basket',
'vegetable basket'}), support=0.0055409336473195734,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'fruit basket'}),
items_add=frozenset({'vegetable basket'}), confidence=0.29411764705882354,
lift=3.081618714249125)]), RelationRecord(items=frozenset({'ginger garlic
paste', 'olive oil'}), support=0.003324560188391744,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'ginger garlic
paste'}), items_add=frozenset({'olive oil'}), confidence=0.21238938053097342,
lift=3.2346813039094875)]), RelationRecord(items=frozenset({'kissan puree',
'paneer'}), support=0.005402410306136584,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'kissan puree'}),
items_add=frozenset({'paneer'}), confidence=0.3786407766990291,
lift=3.8390558525144542)]), RelationRecord(items=frozenset({'pasta', 'maggi'}),
support=0.0034630835295747335,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'pasta'}),
items_add=frozenset({'maggi'}), confidence=0.24752475247524755,
lift=3.456249803071797)]), RelationRecord(items=frozenset({'pasta',
```

```
'mushroom'}), support=0.005402410306136584,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'pasta'}),
items_add=frozenset({'mushroom'}), confidence=0.38613861386138615,
lift=4.873312331233124)]), RelationRecord(items=frozenset({'mushroom', 'pepper
spray'}), support=0.005817980329685552,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'pepper spray'}),
items_add=frozenset({'mushroom'}), confidence=0.3021582733812949,
lift=3.8134275796146295)]), RelationRecord(items=frozenset({'neckrest', 'trolly
bag'}), support=0.0034630835295747335,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'neckrest'}),
items_add=frozenset({'trolly bag'}), confidence=0.2717391304347826,
lift=5.620873302603712)]), RelationRecord(items=frozenset({'whole wheat pasta',
'olive oil'}), support=0.007757307106247403,
ordered_statistics=[OrderedStatistic(items_base=frozenset({'whole wheat
pasta'}), items_add=frozenset({'olive oil'}), confidence=0.26666666666666666,
lift=4.061322081575246)])]
```

1.4.2 Proper Display

[10]:
```
def inspect(results):
    product1         = [tuple(result[2][0][0])[0] for result in results]
    product2         = [tuple(result[2][0][1])[0] for result in results]
    supports    = [result[1] for result in results]
    confidences = [result[2][0][2] for result in results]
    lifts       = [result[2][0][3] for result in results]
    return list(zip(product1, product2, supports, confidences, lifts))
DataFrame_intelligence = pd.DataFrame(inspect(results), columns = ['product1',␣
 ↪'product1', 'Support', 'Confidence', 'Lift'])
```

[11]:
```
DataFrame_intelligence
```

[11]:

	product1	product1	Support	Confidence	Lift
0	almonds	burgers	0.005402	0.265306	3.059497
1	buns	paneer	0.016069	0.324022	3.285277
2	ginger garlic paste	chicken	0.004571	0.292035	4.791372
3	fruit basket	vegetable basket	0.005541	0.294118	3.081619
4	ginger garlic paste	olive oil	0.003325	0.212389	3.234681
5	kissan puree	paneer	0.005402	0.378641	3.839056
6	pasta	maggi	0.003463	0.247525	3.456250
7	pasta	mushroom	0.005402	0.386139	4.873312
8	pepper spray	mushroom	0.005818	0.302158	3.813428
9	neckrest	trolly bag	0.003463	0.271739	5.620873
10	whole wheat pasta	olive oil	0.007757	0.266667	4.061322

[ ]:
```
DataFrame_intelligence.nlargest(n = 10, columns = 'Lift')
```

[ ]:

	Left Hand Side	Right Hand Side	Support	Confidence	Lift
8	neckrest	patanjali honey	0.003463	0.271739	5.620873
3	pasta	escalope	0.005402	0.386139	4.873312
1	light cream	chicken	0.004571	0.292035	4.791372
9	whole wheat pasta	olive oil	0.007757	0.266667	4.061322
5	kissan puree	ground beef	0.005402	0.378641	3.839056
2	mushroom cream sauce	escalope	0.005818	0.302158	3.813428
7	pasta	maggi	0.003463	0.247525	3.456250
4	herb & pepper	ground beef	0.016069	0.324022	3.285277
6	light cream	olive oil	0.003325	0.212389	3.234681
10	pizza cheese	vegetable basket	0.005541	0.294118	3.081619