

EXPERIMENT NO. 9

AIM: Study and Implement Classification based on association rules.

Theory:

Classification based on association rules, often referred to as **Associative Classification**, is a machine learning approach that combines classification and association rule mining to generate accurate and interpretable classifiers. Unlike traditional classifiers, which rely purely on statistical models, associative classification uses frequent item sets and association rules to predict class labels, making it particularly suitable for datasets with meaningful item associations, such as market basket analysis.

The key steps in associative classification include:

1. **Association Rule Mining:** The process begins by identifying frequent item sets from the dataset using algorithms like Apriori or FP-Growth. These item sets are combinations of items that frequently occur together, based on a minimum support threshold.
2. **Rule Generation:** Association rules are then generated from these frequent item sets. A rule typically has the form **IF antecedent THEN consequent**, where the antecedent is a combination of features, and the consequent is the target class label. Each rule is evaluated based on metrics like **support** (the proportion of records containing both antecedent and consequent) and **confidence** (the likelihood that a record containing the antecedent also contains the consequent).
3. **Rule Selection:** Not all generated rules are used in classification. The rules are filtered based on confidence and other criteria to retain only the most predictive and reliable ones.
4. **Classification:** To classify a new instance, the classifier examines the rules that apply to the instance and chooses the rule with the highest confidence or uses a voting mechanism among applicable rules to predict the class label.

By leveraging association rules, this technique provides high interpretability, as the rules generated can be directly analysed and understood, making it easier to extract insights from data. Associative classification has been effectively used in fields where pattern and rule interpretability is crucial, such as healthcare, retail, and finance.

Code:

```
# In[1]:  
import numpy as np  
import matplotlib.pyplot as plt  
import pandas as pd
```

Data Preprocessing

```
# In[2]:
dataset = pd.read_csv('Market_Basket_Optimisation.csv', header = None)
transactions = []
for i in range(0, 7501):
    transactions.append([str(dataset.values[i,j]) for j in range(0, 20)])
```

Training the Apriori model on the dataset

```
# In[3]:
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)
```

Visualising the results

Displaying the first results coming directly from the output of the apriori function

```
#In[4]:
results = list(rules)
```

```
#In[5]:
Results
```

Putting the results well organised into a Pandas DataFrame

```
#In[6]:
def inspect(results):
    lhs      = [tuple(result[2][0][0])[0] for result in results]
    rhs      = [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(lhs, rhs, supports, confidences, lifts))
resultsinDataFrame = pd.DataFrame(inspect(results), columns = ['Left Hand Side', 'Right
Hand Side', 'Support', 'Confidence', 'Lift'])
```

Displaying the results non sorted

```
#In[7]:
resultsinDataFrame
```

Displaying the results sorted by descending lifts

```
#In[8]:
resultsinDataFrame.nlargest(n = 5, columns = 'Lift')
```

Output Snapshots:

```
[RelationRecord(items=frozenset({'chicken', 'light cream'}), support=0.00453272896947073
7, ordered_statistics=[OrderedStatistic(items_base=frozenset({'light cream'}), items_add
=frozenset({'chicken'}), confidence=0.29059829059829057, lift=4.84395061728395)]),
  RelationRecord(items=frozenset({'mushroom cream sauce', 'escalope'}), support=0.0057325
68990801226, ordered_statistics=[OrderedStatistic(items_base=frozenset({'mushroom cream
sauce'}), items_add=frozenset({'escalope'}), confidence=0.3006993006993007, lift=3.79083
2696715049)]),
  RelationRecord(items=frozenset({'pasta', 'escalope'}), support=0.005865884548726837, or
dered_statistics=[OrderedStatistic(items_base=frozenset({'pasta'}), items_add=frozenset
({'escalope'}), confidence=0.3728813559322034, lift=4.700811850163794)]),
  RelationRecord(items=frozenset({'fromage blanc', 'honey'}), support=0.00333288894814024
8, ordered_statistics=[OrderedStatistic(items_base=frozenset({'fromage blanc'}), items_a
dd=frozenset({'honey'}), confidence=0.2450980392156863, lift=5.164270764485569)]),
  RelationRecord(items=frozenset({'ground beef', 'herb & pepper'}), support=0.01599786695
1073192, ordered_statistics=[OrderedStatistic(items_base=frozenset({'herb & pepper'}), i
tems_add=frozenset({'ground beef'}), confidence=0.3234501347708895, lift=3.2919938411349
285)]),
  RelationRecord(items=frozenset({'tomato sauce', 'ground beef'}), support=0.005332622317
024397, ordered_statistics=[OrderedStatistic(items_base=frozenset({'tomato sauce'}), ite
ms_add=frozenset({'ground beef'}), confidence=0.3773584905660377, lift=3.84065948132408
3)]),
  RelationRecord(items=frozenset({'light cream', 'olive oil'}), support=0.003199573390214
638, ordered_statistics=[OrderedStatistic(items_base=frozenset({'light cream'}), items_a
dd=frozenset({'olive oil'}), confidence=0.20512820512820515, lift=3.1147098515519573)]),
  RelationRecord(items=frozenset({'whole wheat pasta', 'olive oil'}), support=0.007998933
475536596, ordered_statistics=[OrderedStatistic(items_base=frozenset({'whole wheat past
a'}), items_add=frozenset({'olive oil'}), confidence=0.2714932126696833, lift=4.12241009
7642296)]),
  RelationRecord(items=frozenset({'pasta', 'shrimp'}), support=0.005065991201173177, orde
red_statistics=[OrderedStatistic(items_base=frozenset({'pasta'}), items_add=frozenset
({'shrimp'}), confidence=0.3220338983050847, lift=4.506672147735896)])]
```

	Left Hand Side	Right Hand Side	Support	Confidence	Lift
0	light cream	chicken	0.004533	0.290598	4.843951
1	mushroom cream sauce	escalope	0.005733	0.300699	3.790833
2	pasta	escalope	0.005866	0.372881	4.700812
3	fromage blanc	honey	0.003333	0.245098	5.164271
4	herb & pepper	ground beef	0.015998	0.323450	3.291994
5	tomato sauce	ground beef	0.005333	0.377358	3.840659
6	light cream	olive oil	0.003200	0.205128	3.114710
7	whole wheat pasta	olive oil	0.007999	0.271493	4.122410
8	pasta	shrimp	0.005066	0.322034	4.506672

	Left Hand Side	Right Hand Side	Support	Confidence	Lift
3	fromage blanc	honey	0.003333	0.245098	5.164271
0	light cream	chicken	0.004533	0.290598	4.843951
2	pasta	escalope	0.005866	0.372881	4.700812
8	pasta	shrimp	0.005066	0.322034	4.506672
7	whole wheat pasta	olive oil	0.007999	0.271493	4.122410

Learning Outcome: