

BÀI THỰC HÀNH SỐ 6-7

Luật kết hợp

Một số lệnh sử dụng khi khai phá luật kết hợp

#các thư viện sử dụng để dùng trong khai phá luật kết hợp

```
import numpy as np
```

```
import matplotlib.pyplot as plt
```

```
import pandas as pd
```

```
from apyori import apriori
```

```
#đọc dữ liệu
```

```
data = pd.read_csv('d:\\hoc python\\105.csv', header=None)
```

#hiển thị 10 bảng ghi đầu tiên

```
data.head(10)
```

#Khởi tạo biến ghi để đọc dữ liệu

```
records = []
```

#đọc dữ liệu vào mảng

```
for i in range(0,90):
```

```
    records.append([str(data.values[i,j]) for j in range(0,10)])
```

#khai phá luật kết hợp với 10 thuộc tính, độ hỗ trợ 0,01, độ tin cậy 0,3

```
association_rules = apriori(records, min_support=0.01, min_confidence=0.3,  
min_lift=3, min_length=2)
```

```
association_results = list(association_rules)
```

```
#in toàn bộ luật
```

```
print(association_results)
```

#in từng luật

for item in association_results:

first index of the inner list

Contains base item and add item

pair = item[0]

items = [x for x in pair]

print("Rule: " + items[0] + " -> " + items[1])

#second index of the inner list

print("Support: " + str(item[1]))

#third index of the list located at 0th

#of the third index of the inner list

print("Confidence: " + str(item[2][0][2]))

print("Lift: " + str(item[2][0][3]))

print("=====")

Bài 1: Làm theo bài mẫu **chuyendoidl_luatkethop** sau đó làm thực hiện trên tập dữ liệu sau:

dataset = [['Milk', 'Onion', 'Apple', 'Kidney Beans', 'Eggs', 'Yogurt'],

['Banana', 'Onion', 'Apple', 'Kidney Beans', 'Eggs', 'Yogurt'],

['Milk', 'Apple', 'Kidney Beans', 'Eggs'],

['Milk', 'Ice cream', 'Corn', 'Kidney Beans', 'Yogurt'],

['Corn', 'Apple', 'Onion', 'Kidney Beans', 'Ice cream', 'Eggs']]

Sau đó thay đổi tìm các tập mục có độ hỗ trợ 60%, 80%, tìm luật cho các tập mục có độ hỗ trợ tương ứng và lọc những luật có độ tin cậy 85%,90% cho từng mỗi mức độ hỗ trợ. So sánh và rút ra nhận xét.

Bài 2: Thực hiện khai phá luật kết hợp cho tệp tin 105OK.csv đã được thực hiện trên bài thực hành 3,4 với độ hỗ trợ 40% và độ tin cậy 60%, thử với các mức khác và đưa ra nhận xét.

Bài 3: Cho dữ liệu giao dịch trong file Market_Basket_AR, đọc và chuyển đổi dữ liệu về dạng 0, 1. Thực hiện khai phá luật kết hợp với 3 mức độ hỗ trợ và tin cậy khác nhau.

Bài 4: Tìm luật kết hợp với thuật toán fp-Growth với dữ liệu và yêu cầu như bài 2,3.