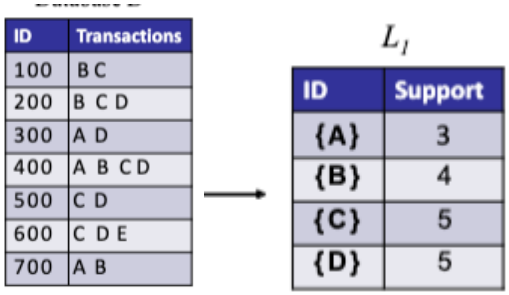
Course: 634
Subject: Midterm Project
Topic: Apriori Algorithm Implementation
Name: Sheethal Mathew
Email:ssm29@njit.edu

Apriori Algorithm Implementation

Given: Minimum support =2, Minimum Confidence=50%



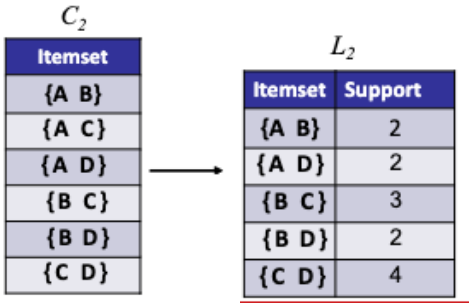
ID	Transactions
100	B C
200	B C D
300	A D
400	A B C D
500	C D
600	C D E
700	A B

L_1

ID	Support
{A}	3
{B}	4
{C}	5
{D}	5

Figure1

K=2



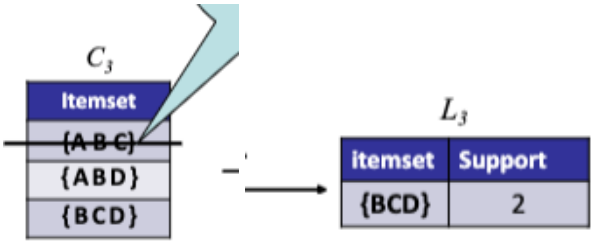
C_2

Itemset
{A B}
{A C}
{A D}
{B C}
{B D}
{C D}

L_2

Itemset	Support
{A B}	2
{A D}	2
{B C}	3
{B D}	2
{C D}	4

Figure2
K=3



C_3

Itemset
{A B C}
{ABD}
{BCD}

L_3

Itemset	Support
{BCD}	2

Figure 1, Figure 2, Figure 3 from Page.14, "Module_2_Association_Rule_Mining.pdf," Yasser Abduallah, NJIT

Contents

Introduction

Apriori is an algorithm appropriate in using to solve problems that require computing frequency of transaction. Examples are every time item milk is bought at grocers, the itemset frequency count increases. Then calculate the support, and extract the new itemset after eliminating itemset that did not meet minimum support.

Implementation Overview

The implementation of apriori does not use the inbuilt apriori library in python.
The data is read from csv dataset. When running the program you have to provide support and confidence. Support and Confidence is a floating point number between 0 and 1 to indicate a percentage.
It is ran on Pycharm with Pyenv environment.

Assumption

Assumption made is that the dataset provided has no newline.

Requirement

Software

Python3, Pycharm, Pyenv environment are required to run the program.
Instead of JupyterLab Pycharm was used

Hardware

MacBook Pro, Apple M1chip, MacOS 12.0.1

List of Dataset Files

Each dataset is 20 transactions
dataset_test.csv
dataset_kmart.csv
dataset_amazon.csv
dataset_bestbuy.csv
dataset_nike.csv

List of source code Files

Apriori.py → It takes dataset as csv files, minimum support, minimum confidence percentage as parameters. Calculate the itemset when k=1, that will meet the minimum support, and get the new Lset. Increase the counter for k by 1. Now you have C2, candidate set2 that meets the criteria and get L2 after meeting min support again.

How to run the program

Download Pycharm from https://www.jetbrains.com/pycharm/
Click on view->tool windows -> terminal
Enter the following without the quotes.

“Python3 apriori.py -f dataset_test.csv -s 0.1 -c 0.6”

```
(venv) sheethalmathew@Sheethals-MBP pythonProject1 % python3 aprioriFinal.py -f dataset_test.csv -s 0.1 -c 0.6
aprioriFinal.py:142: DeprecationWarning: 'U' mode is deprecated
  with open(fname, "rU") as file_iter:

k1
 LSet

item: ('Laptop"',) , 0.182
item: ('"Tablet',) , 0.182
item: ('Charger', '"Headphone') , 0.182
item: ('Laptop"', '"Headphone') , 0.182
item: ('Headphone"', 'Laptop') , 0.182
```

Figure1: Apriori Execution

Result

I used the following to run the program: “python apriori.py -f dataset_test.csv -s 0.1 -c 0.6”
In this the minimum support I give is 0.1 and minimum confidence is 0.6
The results show item set that meets the minimum support

```
k1
 LSet

item: ('"Tablet',) , 0.182
item: ('Laptop"',) , 0.182
item: ('Headphone"', 'Laptop') , 0.182
item: ('"Headphone', 'Charger') , 0.182
item: ('"Headphone', 'Laptop"') , 0.182
item: ('Laptop',) , 0.273
item: ('Headphone"',) , 0.273
item: ('Powerbank"',) , 0.273
item: ('Charger',) , 0.364
item: ('"Headphone',) , 0.455

----------------------- RULES:
Rule: ('Headphone"',) ==> ('Laptop',) , 0.667
Rule: ('Laptop',) ==> ('Headphone"',) , 0.667
Rule: ('Laptop"',) ==> ('"Headphone',) , 1.000
```

Figure2: Apriori on Stationary items

Ex. Set of tablet =0.182 support
0.182>0.1, therefore include it in the itemset

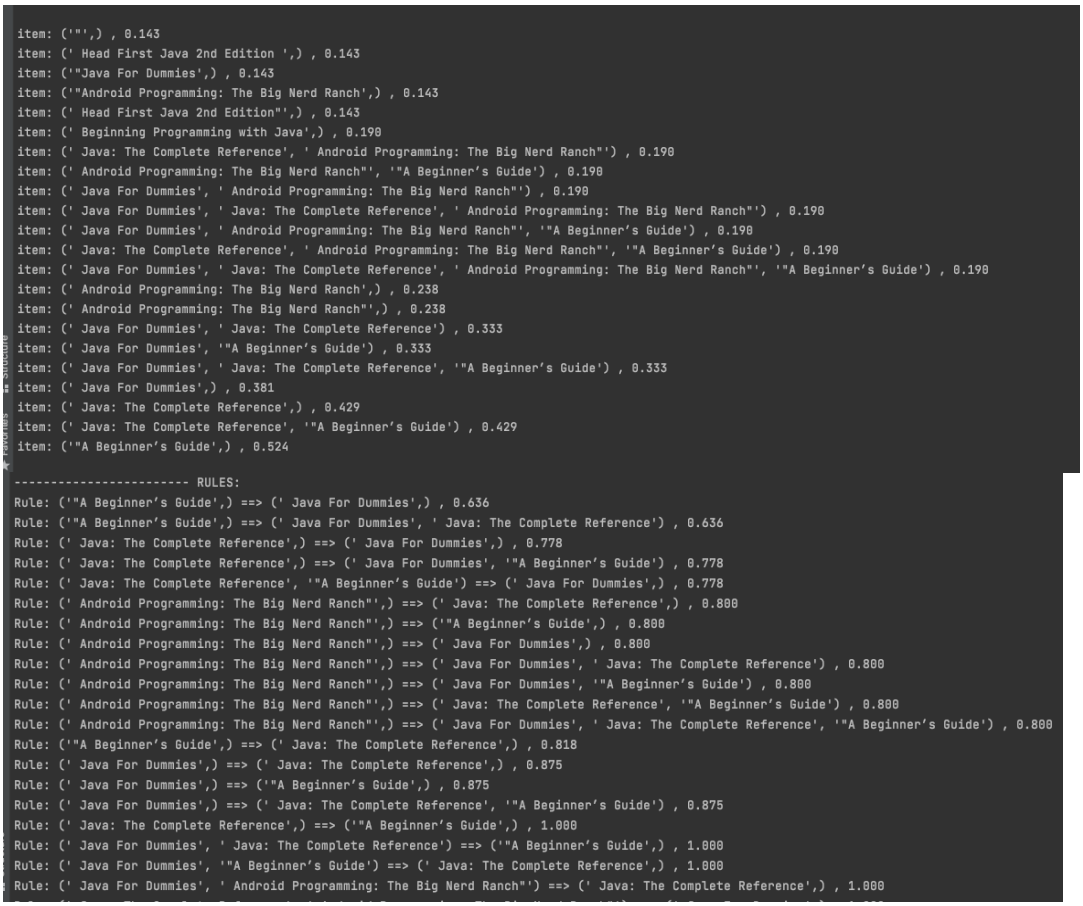


Figure3: Amazon Itemset and condfidence after runningApriori

```
Rule: (' Lab Top',) ==> (' Lab Top Case',) , 0.600
Rule: (' Lab Top',) ==> (' Anti-Virus',) , 0.600
Rule: (' Speakers"',) ==> (' External Hard-Drive',) , 0.600
Rule: (' Printer', ' Microsoft Office') ==> (' Speakers',) , 0.600
Rule: (' Speakers"',) ==> (' Lab Top Case', '"Digital Camera ') , 0.600
Rule: (' Anti-Virus', ' Microsoft Office') ==> (' External Hard-Drive"',) , 0.600
Rule: (' Speakers"',) ==> (' Anti-Virus', ' Lab Top Case') , 0.600
Rule: (' Anti-Virus', ' Microsoft Office') ==> (' Printer',) , 0.600
Rule: (' Printer', ' Microsoft Office') ==> (' Anti-Virus',) , 0.600
Rule: (' Speakers"',) ==> (' Lab Top Case', ' External Hard-Drive') , 0.600
Rule: ('"Lab Top', ' Flash Drive') ==> (' Lab Top Case',) , 0.600
Rule: (' Speakers"',) ==> (' Anti-Virus', ' External Hard-Drive') , 0.600
Rule: (' Lab Top',) ==> (' Lab Top Case', '"Digital Camera ') , 0.600
Rule: (' Lab Top', '"Digital Camera ') ==> (' Lab Top Case',) , 0.600
Rule: (' Lab Top',) ==> (' Anti-Virus', '"Digital Camera ') , 0.600
Rule: (' Lab Top', '"Digital Camera ') ==> (' Anti-Virus',) , 0.600
Rule: ('"Lab Top', ' Flash Drive') ==> (' Printer',) , 0.600
Rule: (' Anti-Virus', ' Microsoft Office') ==> (' Flash Drive', ' External Hard-Drive"') , 0.600
Rule: (' Anti-Virus', ' Flash Drive', ' Microsoft Office') ==> (' External Hard-Drive"',) , 0.600
Rule: (' Anti-Virus', ' Lab Top Case', ' Flash Drive') ==> (' External Hard-Drive"',) , 0.600
Rule: (' Anti-Virus', ' Lab Top Case', ' Flash Drive') ==> (' Printer',) , 0.600
item: (' Microsoft Office"',) , 0.143
item: (' Desk Top',) , 0.143
item: (' External Hard-Drive"',) , 0.143
item: ('"Desk Top',) , 0.143
item: (' Lab Top Case', ' External Hard-Drive"') , 0.143
item: (' Anti-Virus', '"Digital Camera ') , 0.143
item: (' Anti-Virus', ' Speakers"') , 0.143
item: (' Printer', '"Digital Camera ') , 0.143
item: (' Microsoft Office', ' External Hard-Drive"') , 0.143
item: (' Lab Top', ' Lab Top Case') , 0.143
item: ('"Desk Top', ' Printer') , 0.143
item: ('"Lab Top', ' External Hard-Drive') , 0.143
item: (' Anti-Virus', ' Lab Top') , 0.143
item: (' Speakers', ' Anti-Virus"') , 0.143
item: (' Speakers"', ' External Hard-Drive') , 0.143
item: (' Flash Drive', ' External Hard-Drive') , 0.143
item: (' Anti-Virus', ' External Hard-Drive"') , 0.143
item: (' Anti-Virus', ' Speakers') , 0.143
item: ('"Lab Top', ' Printer') , 0.143
item: (' Microsoft Office"', ' Flash Drive') , 0.143
item: (' Microsoft Office"', ' Printer') , 0.143
```

Figure4: Bestbuy itemset after apriori

```
item: (' Bed Skirts', ' Kids Bedding', ' Sheets "') , 0.143
item: (' Sheets "', ' Bed Skirts', ' Shams') , 0.143
item: (' Kids Bedding', ' Bed Skirts', ' Shams') , 0.143
item: (' Kids Bedding', ' Sheets "', ' Shams') , 0.143
item: (' Sheets "', ' Kids Bedding', ' Bed Skirts', ' Shams') , 0.143
item: (' Kids Bedding', ' Shams') , 0.190
item: (' Bedspreads', ' Bed Skirts') , 0.190
item: (' Bedspreads', ' Shams') , 0.190
item: (' Bedspreads', ' Bed Skirts', ' Shams') , 0.190
item: (' Bedspreads',) , 0.238
item: (' Sheets',) , 0.238
item: (' Bed Skirts', ' Shams') , 0.286
item: (' Kids Bedding', ' Bed Skirts') , 0.286
```

```
Rule: (' Sheets',) ==> (' Bed Skirts',) , 0.600
Rule: (' Bedspreads',) ==> (' Sheets',) , 0.600
Rule: (' Sheets',) ==> (' Bedspreads',) , 0.600
Rule: (' Bed Skirts',) ==> (' Shams',) , 0.667
Rule: (' Bed Skirts',) ==> (' Kids Bedding',) , 0.667
Rule: (' Bed Skirts', ' Shams') ==> (' Bedspreads',) , 0.667
Rule: (' Shams',) ==> (' Bed Skirts',) , 0.750
Rule: (' Kids Bedding', ' Shams') ==> (' Bed Skirts',) , 0.750
Rule: (' Kids Bedding', ' Shams') ==> (' Sheets "',) , 0.750
Rule: (' Kids Bedding', ' Shams') ==> (' Bed Skirts', ' Sheets "') , 0.750
Rule: (' Bedspreads',) ==> (' Bed Skirts',) , 0.800
Rule: (' Bedspreads',) ==> (' Shams',) , 0.800
Rule: (' Bedspreads',) ==> (' Bed Skirts', ' Shams') , 0.800
Rule: (' Kids Bedding',) ==> (' Bed Skirts',) , 0.857
Rule: (' Sheets "',) ==> (' Kids Bedding',) , 1.000
Rule: (' Quilts"',) ==> ('"Decorative Pillows',) , 1.000
Rule: (' Sheets "',) ==> (' Bed Skirts',) , 1.000
Rule: (' Sheets "',) ==> (' Shams',) , 1.000
Rule: (' Quilts',) ==> ('"Decorative Pillows',) , 1.000
Rule: (' Sheets "',) ==> (' Kids Bedding', ' Bed Skirts') , 1.000
Rule: (' Sheets "', ' Bed Skirts') ==> (' Kids Bedding',) , 1.000
Rule: (' Kids Bedding', ' Sheets "') ==> (' Bed Skirts',) , 1.000
Rule: (' Sheets "',) ==> (' Bed Skirts', ' Shams') , 1.000
Rule: (' Bed Skirts', ' Sheets "') ==> (' Shams',) , 1.000
Rule: (' Sheets "', ' Shams') ==> (' Bed Skirts',) , 1.000
Rule: (' Bedspreads', ' Bed Skirts') ==> (' Shams',) , 1.000
```

```
item: (' Socks', ' Rash Guard', ' Modern Pants', '"Running Shoe') , 0.238
item: (' Rash Guard', ' Modern Pants', ' Tech Pants', ' Sweatshirts') , 0.238
item: (' Modern Pants', '"Running Shoe', ' Socks', ' Rash Guard', ' Sweatshirts') , 0.238
item: (' Swimming Shirt',) , 0.286
item: (' Hoodies',) , 0.286
item: ('"Running Shoe', ' Rash Guard') , 0.286
item: (' Rash Guard', ' Modern Pants') , 0.286
item: (' Rash Guard', ' Tech Pants') , 0.286
item: (' Rash Guard', ' Sweatshirts') , 0.286
item: (' Socks', ' Modern Pants') , 0.286
item: ('"Running Shoe', ' Swimming Shirt') , 0.286
item: ('"Running Shoe', ' Modern Pants') , 0.286
item: ('"Running Shoe', ' Modern Pants', ' Sweatshirts') , 0.286
item: (' Sweatshirts', ' Rash Guard', ' Modern Pants') , 0.286
item: ('"Running Shoe', ' Socks', ' Modern Pants') , 0.286
item: (' Socks', ' Modern Pants', ' Sweatshirts') , 0.286
item: (' Socks', ' Modern Pants', '"Running Shoe', ' Sweatshirts') , 0.286
item: (' Tech Pants',) , 0.333
item: (' Modern Pants',) , 0.333
item: (' Sweatshirts', ' Modern Pants') , 0.333
item: (' Socks', ' Sweatshirts') , 0.381
item: ('"Running Shoe', ' Socks', ' Sweatshirts') , 0.381
item: ('"Running Shoe', ' Sweatshirts') , 0.429
item: (' Rash Guard',) , 0.476
item: (' Sweatshirts',) , 0.476
item: ('"Running Shoe', ' Socks') , 0.476
```

Figure5: Kmart dataset after running Apriori

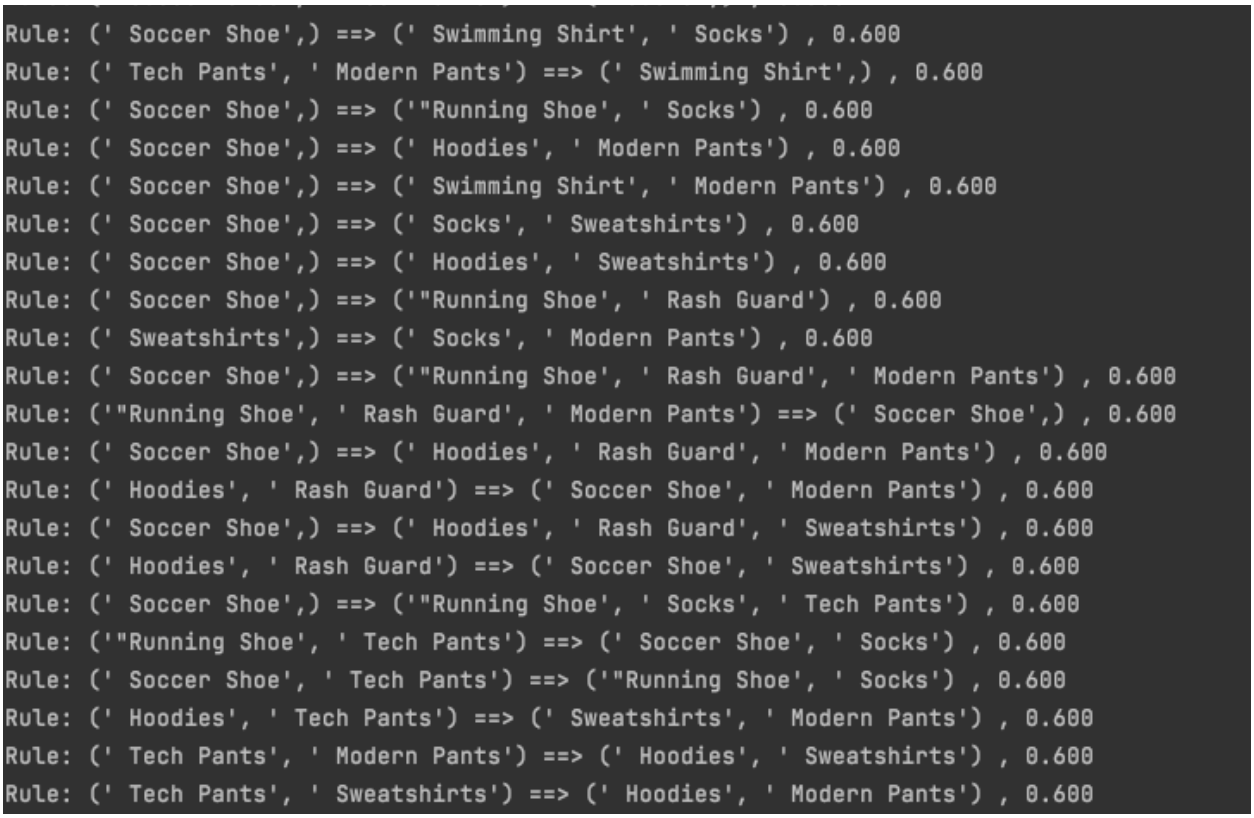


Figure5: Nik dataset after running Apriori

Logs and Execution Result

1 dataset result for dataset_amazon.csv log is provided

```
(venv) sheethalmathew@Sheethals-MBP pythonProject1 % clear
(venv) sheethalmathew@Sheethals-MBP pythonProject1 % python3 aprioriFinal.py -f dataset_amazon.csv -s 0.1
-c 0.6
aprioriFinal.py:142: DeprecationWarning: 'U' mode is deprecated
  with open(fname, "rU") as file_iter:

k1
 LSet

k 3
k 4
k 5
k 6
item: (' Head First Java 2nd Edition ',) , 0.143
item: (' Head First Java 2nd Edition"',) , 0.143
item: ('"',) , 0.143
item: ('"Java For Dummies',) , 0.143
item: ('"Android Programming: The Big Nerd Ranch',) , 0.143
item: (' Beginning Programming with Java',) , 0.190
item: ('"A Beginner’s Guide', ' Android Programming: The Big Nerd Ranch"') , 0.190
item: (' Java: The Complete Reference', ' Android Programming: The Big Nerd Ranch"') , 0.190
item: (' Android Programming: The Big Nerd Ranch"', ' Java For Dummies') , 0.190
item: ('"A Beginner’s Guide', ' Android Programming: The Big Nerd Ranch"', ' Java For Dummies') , 0.190
item: (' Java: The Complete Reference', ' Android Programming: The Big Nerd Ranch"', ' Java For Dummies') ,
0.190
item: ('"A Beginner’s Guide', ' Android Programming: The Big Nerd Ranch"', ' Java: The Complete Reference') ,
0.190
item: ('"A Beginner’s Guide', ' Java: The Complete Reference', ' Android Programming: The Big Nerd Ranch"', '
Java For Dummies') , 0.190
item: (' Android Programming: The Big Nerd Ranch',) , 0.238
item: (' Android Programming: The Big Nerd Ranch"',) , 0.238
item: ('"A Beginner’s Guide', ' Java For Dummies') , 0.333
item: (' Java: The Complete Reference', ' Java For Dummies') , 0.333
item: ('"A Beginner’s Guide', ' Java: The Complete Reference', ' Java For Dummies') , 0.333
item: (' Java For Dummies',) , 0.381
item: (' Java: The Complete Reference',) , 0.429
item: ('"A Beginner’s Guide', ' Java: The Complete Reference') , 0.429
item: ('"A Beginner’s Guide',) , 0.524

---------- RULES:
Rule: ('"A Beginner’s Guide',) ==> (' Java For Dummies',) , 0.636
Rule: ('"A Beginner’s Guide',) ==> (' Java: The Complete Reference', ' Java For Dummies') , 0.636
Rule: (' Java: The Complete Reference',) ==> (' Java For Dummies',) , 0.778
Rule: (' Java: The Complete Reference',) ==> ('"A Beginner’s Guide', ' Java For Dummies') , 0.778
Rule: ('"A Beginner’s Guide', ' Java: The Complete Reference') ==> (' Java For Dummies',) , 0.778
Rule: (' Android Programming: The Big Nerd Ranch"',) ==> ('"A Beginner’s Guide',) , 0.800
Rule: (' Android Programming: The Big Nerd Ranch"',) ==> (' Java: The Complete Reference',) , 0.800
Rule: (' Android Programming: The Big Nerd Ranch"',) ==> (' Java For Dummies',) , 0.800
Rule: (' Android Programming: The Big Nerd Ranch"',) ==> ('"A Beginner’s Guide', ' Java For Dummies') , 0.800
Rule: (' Android Programming: The Big Nerd Ranch"',) ==> (' Java: The Complete Reference', ' Java For Dummies')
, 0.800
Rule: (' Android Programming: The Big Nerd Ranch"',) ==> ('"A Beginner’s Guide', ' Java: The Complete
Reference') , 0.800
Rule: (' Android Programming: The Big Nerd Ranch"',) ==> (' Java For Dummies', '"A Beginner’s Guide', ' Java: The
Complete Reference') , 0.800
Rule: ('"A Beginner’s Guide',) ==> (' Java: The Complete Reference',) , 0.818
Rule: (' Java For Dummies',) ==> ('"A Beginner’s Guide',) , 0.875
Rule: (' Java For Dummies',) ==> (' Java: The Complete Reference',) , 0.875
Rule: (' Java For Dummies',) ==> ('"A Beginner’s Guide', ' Java: The Complete Reference') , 0.875
Rule: (' Java: The Complete Reference',) ==> ('"A Beginner’s Guide',) , 1.000
Rule: ('"A Beginner’s Guide', ' Java For Dummies') ==> (' Java: The Complete Reference',) , 1.000
Rule: (' Java: The Complete Reference', ' Java For Dummies') ==> ('"A Beginner’s Guide',) , 1.000
Rule: ('"A Beginner’s Guide', ' Android Programming: The Big Nerd Ranch"') ==> (' Java For Dummies',) , 1.000
Rule: (' Android Programming: The Big Nerd Ranch"', ' Java For Dummies') ==> ('"A Beginner’s Guide',) , 1.000
Rule: (' Java: The Complete Reference', ' Android Programming: The Big Nerd Ranch"') ==> (' Java For Dummies',)
, 1.000
Rule: (' Android Programming: The Big Nerd Ranch"', ' Java For Dummies') ==> (' Java: The Complete Reference',)
, 1.000
Rule: ('"A Beginner’s Guide', ' Android Programming: The Big Nerd Ranch"') ==> (' Java: The Complete
Reference',) , 1.000
Rule: (' Java: The Complete Reference', ' Android Programming: The Big Nerd Ranch"') ==> ('"A Beginner’s
Guide',) , 1.000
```

References

Yasser Abduallah, “Module_2_Association_Rule_Mining.pdf,” Pg.14, NJIT

http://www.vucreations.com/