ΕΡΓΑΣΙΑ 4 – Εξόρυξη Κανόνων Συσχέτισης

1. Apriori με το χέρι

Εκτελέστε τον αλγόριθμο Apriori πάνω στα δεδομένα φαστφουντάδικου με όρια υποστήριξης s=33.34% και εμπιστοσύνης c=60%. Δείξτε τα υποψήφια και τα συχνά στοιχειοσύνολα σε κάθε πέρασμα της βάσης. Δώστε τη λίστα με τα τελικά συχνά στοιχειοσύνολα και τη λίστα με τους κανόνες συσχέτισης που παράγονται από αυτά ταξινομημένους ως προς την εμπιστοσύνη. Για τους τελικούς κανόνες συσχέτισης (δηλαδή για αυτούς που έχουν c>=60%) υπολογίστε και το lift.

|  |  |
| --- | --- |
| **Transaction ID** | **Items** |
| T1 | Γύρος, Ψωμάκι, Μουστάρδα |
| T2 | Γύρος, Ψωμάκι |
| T3 | Γύρος, Πίτα, Πατάτες |
| T4 | Πατάτες, Πίτα |
| T5 | Πατάτες, Μουστάρδα |
| T6 | Γύρος, Πίτα, Πατάτες, Μουστάρδα |

# Απάντηση:

Όταν γράφει ότι το support είναι 33,34% σημαίνει ότι πρέπει να έχει σαν όριο τουλάχιστον 2 συναλλαγές.

|  |  |  |
| --- | --- | --- |
| Pass(k) | Υποψήφια k-itemsets και η υποστήριξή τους | Συχνότητα k-itemsets |
| K=1 | Γύρος(4), Ψωμάκι(2), Μουστάρδα(3), Πίτα(3), Πατάτες(4) | Γύρος, Ψωμάκι, Μουστάρδα, Πίτα, Πατάτες |
| K=2 | {Γύρος, Ψωμάκι}(2), {Γύρος, Μουστάρδα}(2), {Γύρος, Πίτα}(2), {Γύρος, Πατάτες}(2), {Ψωμάκι, Μουστάρδα}(1), ~~{Ψωμάκι, Πίτα}(~~0), ~~{Ψωμάκι, Πατάτες}(~~0), {Μουσταρδα, Πίτα}(1), {Μουστάρδα, Πατάτες}(2), {Πίτα, Πατάτες}(3) | {Γύρος, Ψωμάκι},  {Γύρος,Μουστάρδα},  {Γύρος, Πίτα}, {Γύρος, Πατάτες},  {Μουστάρδα,Πατάτες},  {Πίτα, Πατάτες} |
| K=3 | {Γύρος, Πίτα, Πατάτες}(2), {Γύρος, Μουστάρδα, Πίτα}(1), {Γύρος, Μουστάρδα, Πατάτες}(1) | {Γύρος, Πίτα, Πατάτες} |
| K=4 | {Γύρος, Μουστάρδα, Πίτα, Πατατες}(1) |  |

Όλα τα πιο συχνά σύνολα: {Γύρος}, {Ψωμάκι}, {Μουστάρδα}, {Πίτα}, {Πατάτες}, {Γύρος, Ψωμάκι},{Γύρος, Μουστάρδα}, {Γύρος, Πίτα}, {Γύρος, Πατάτες},{Μουστάρδα, Πατάτες}, {Πίτα, Πατάτες}, {Γύρος, Πίτα, Πατάτες}.

Association rules:

{Γύρος, Ψωμάκι} δημιουργεί: Γύρος 🡺 Ψωμάκι (2/6=0.33, 2/4=0.5) και

**Ψωμάκι** 🡺**Γύρος (2/6=0.33, 2/2=1);**

{Γύρος, Μουστάρδα} δημιουργεί: Γύρος 🡺 Μουστάρδα (2/6=0.33,2/4=0.5) και

**Μουστάρδα 🡺 Γύρος (2/6=0.33, 2/3=0.75)**

{Γύρος, Πίτα} δημιουργεί: Γύρος 🡺 Πίτα (0.33, 0.5) και

**Πίτα** 🡺**Γύρος (2/6=0.33, 2/3=0.66);**

{Γύρος, Πατάτες} δημιουργεί: Γύρος 🡺 Πατάτες (0.33, 0.5) και

Πατάτες 🡺Γύρος (2/6=0.33, 2/4=0.5);

{Μουστάρδα, Πατάτες} δημιουργεί: **Μουστάρδα🡺 Πατάτες(3/6=0.5, 2/3=0.75)**

Πατάτες🡺 Μουστάρδα(3/6=0.5, 2/4=0.5)

{Πίτα, Πατάτες} δημιουργεί: **Πίτα** 🡺**Πατάτες (3/6=0.5, 3/3=1)** και

**Πατάτες** 🡺**Πίτα (3/6=0.5, 3/4=0.75);**

{Γύρος, Πίτα, Πατάτες} δημιουργεί: Γύρος 🡺Πίτα ^ Πατάτες (2/6=0.33, 2/4=0.5),

**Πίτα** 🡺**Πατάτες ^ Γύρος (2/6=0.33, 2/3=0.66)**, **lift=2/8=25%**

Πατάτες 🡺Πίτα ^ Γύρος (2/6=0.33, 2/4=0.5), **Γύρος ^ Πίτα** 🡺**Πατάτες(2/6=0.33, 2/2=1) lift=2/8=25%**, **Γύρος ^ Πατάτες** 🡺 **Πίτα(2/6=0.33, 2/2=1) lift=2/6=33,3%** και **Πίτα ^ Πατάτες** 🡺 **Γύρος(2/6=0.33, 2/3=0.66) lift=2/12=16.6%**.

1. Apriori με το WEKA

Μετασχηματίστε κατάλληλα το προηγούμενο dataset και τρέξτε τον apriori στο WEKA.

Απάντηση: 