Assignment 1

1. You are given a transaction data shown below from a fast food restaurant. For simplicity, we assign the meal items short names [M1-M5]. For all the min\_sup=2/9 and min\_conf=7/9. Apply Apriori and identify all k-frequent itemsets. Find all the strong association rules and note their confidence.

|  |  |
| --- | --- |
| Meal Item | List of Items |
| Order 1 | {M1, M2, M5} |
| Order 2 | {M2, M4} |
| Order 3 | {M2, M3} |
| Order 4 | {M1, M2, M4} |
| Order 5 | {M1, M3} |
| Order 6 | {M2, M3} |
| Order 7 | {M1, M3} |
| Order 8 | {M1, M2, M3, M5} |
| Order 9 | {M1, M2, M3} |

1. Define maximal and closed frequent itemset. Identify the above from the database:

|  |  |
| --- | --- |
| Transaction ID | Items |
| T1 | {A, C, T, W} |
| T2 | {C, D, W} |
| T3 | {A, C, T, W} |
| T4 | {A, C, D, W} |
| T5 | {A, C, D, T, W} |
| T6 | { C, D, T } |

1. Consider the database d shown in the table below. Consider min\_sup =60% and min\_conf=80%. Apply Apriori and identify all k-frequent itemsets. Find all the strong association rules and note their confidence.

|  |  |
| --- | --- |
| TID | Items |
| T100 | {M, O, N, K, E, Y} |
| T200 | {D, O, N, K, E, Y } |
| T300 | {M, A, K, E} |
| T400 | {M, U, C, K, Y} |
| T500 | {C, O, O, K, I ,E} |

1. Consider the transaction database as follows and indicate closed and maximal frequent item sets

|  |  |
| --- | --- |
| TID | Items |
| 1 | {A, B, C} |
| 2 | {A, B, C, D} |
| 3 | {B, C, E} |
| 4 | {A, C, D, E} |
| 5 | {D, E} |

1. Draw the decision tree for the following dataset:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Color | Type | Doors | Tires | Class |
| Red | SUV | 2 | Whitewall | + |
| Blue | Minivan | 4 | Whitewall | - |
| Green | Car | 4 | Whitewall | - |
| Red | Minivan | 4 | Blackwall | - |
| Green | Car | 2 | Blackwall | + |
| Green | SUV | 4 | Blackwall | - |
| Blue | SUV | 2 | Blackwall | - |
| Blue | Car | 2 | Whitewall | + |
| Red | SUV | 2 | Blackwall | - |
| Blue | Car | 4 | Blackwall | - |

1. Construct a decision tree for the following data:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| age | income | student | credit\_rating | buys\_computer |
| <=30 | high | no | fair | no |
| <=30 | high | no | excellent | no |
| 31…40 | high | no | fair | yes |
| >40 | medium | no | fair | yes |
| >40 | low | yes | fair | yes |
| >40 | low | yes | excellent | no |
| 31…40 | low | yes | excellent | yes |
| <=30 | medium | no | fair | no |
| <=30 | low | yes | fair | yes |
| >40 | medium | yes | fair | yes |
| <=30 | medium | yes | excellent | yes |
| 31…40 | medium | no | excellent | yes |
| 31…40 | high | yes | fair | yes |
| >40 | medium | no | excellent | no |