Data Science Hw 4

工科海洋四 B07505015 梁瑞翔

Problem 1

(a)

TID	Items
1	ACDE
2	CDE
3	AC
4	BE
5	ABDE
6	ABCDE

Itemset	Sup
{A}	4
{B}	3
{C}	4
{D}	4
{E}	5

Itemset	Sup
{A}	4
{B}	3
{C}	4
{D}	4
{E}	5

Itemset	Sup
{A B}	2
{A C}	3
{A D}	3
{A E}	3
{B C}	1
{B D}	2
{B E}	3
{C D}	3
{C E}	3
{D E}	4

Itemset	Sup
{A C}	3
{A D}	3
{A E}	3
{B E}	3
{C D}	3
{C E}	3
{D E}	4

Itemset	Sup
{A B E}	2
{A C D}	2
{A C E}	2
{A D E}	3
{B C E}	1
{B D E}	2
{C D E}	3

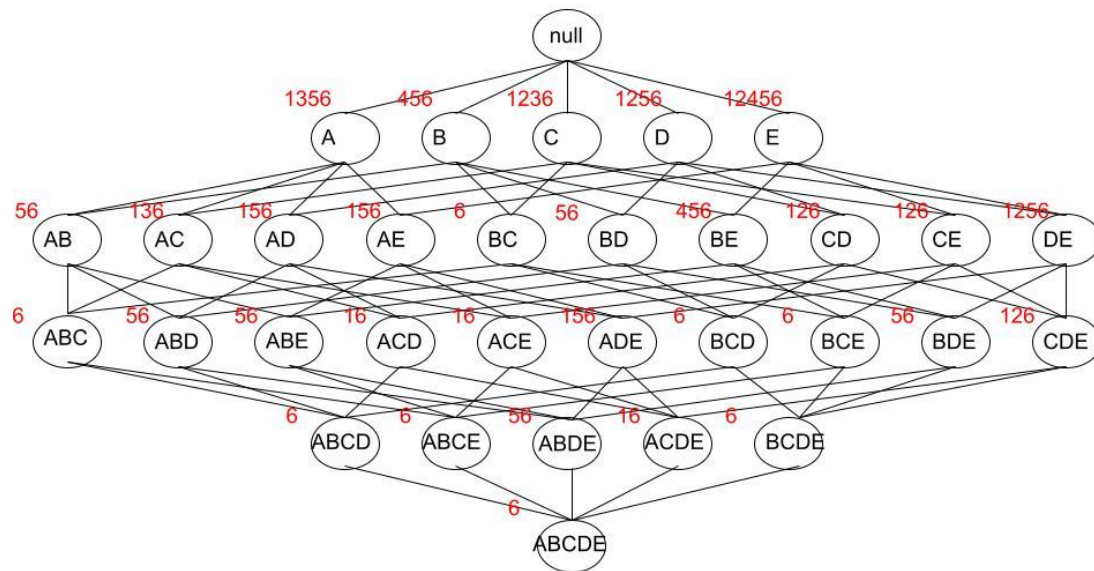| Itemset	Sup
{A D E}	3
{C D E}	3

Itemset	Sup
{A C D E}	2

Itemset	Sup

(b)

{A} {B} {C} {D} {E} {A C} {A D} {A E} {B E} {C D} {C E} {D E} {A D E} {C D E}

(c)

(d)

{A}->{D} s=0.5, c=0.75

{A}->{E} s=0.5, c=0.75

{A}->{DE} s=0.5, c=0.75

{B}->{E} s=0.5, c=1

{C}->{D} s=0.5, c=0.75

(e)

{B}->{D} s=0.33, c=0.66, l=0.66/0.75<1

(a)

Dataset T1 has multiple frequent 1-itemsets overlapping, indicating that it will produce the longest frequent itemset(T3 does not meet the support threshold)

(b)

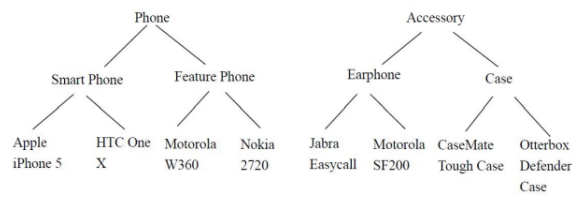T1: 2(ABCDE FGHIJ)

T2: 1(ABG)

T3: 3(ABC CD HIJ)

(c)

T1: 2(ABCDE FGHIJ)

T2: 2(AB ABG)

T3: 6(C D ABC BC CD HIJ)

Problem 3

Phone
- Smart Phone
  - Apple iPhone 5
  - HTC One X
- Feature Phone
  - Motorola W360
  - Nokia 2720

Accessory
- Earphone
  - Jabra Easycall
  - Motorola SF200
- Case
  - CaseMate Tough Case
  - Otterbox Defender Case

Tx	Items bought
100	Apple iPhone 5, Jabra Easycall, CaseMate Tough Case
200	Motorola W360, Motorla SF200, CaseMate Tough Case
300	Nokia2720, Jabra Easycall
400	HTC One X, Jabra Easycall
500	Motorola W360, Motorla SF200, HTC One X
600	Apple iPhone 5, Otterbox Defender Case
700	HTC One X, Otterbox Defender Case
800	Nokia2720
900	Apple iPhone 5, CaseMate Tough Case
1000	Motorola W360, Motorola SF200

10

(a)

Earphone->Phone

$C=0.6/P(Earphone)=1$

$S=6/10=0.6$

(b)

Case->Earphone

$C=0.2/P(Case)=0.4$

$S=2/10=0.2$