Abdul Azzam Ajhari
2201921962

Table 1 shows a list of transactions happening in supermarket “ABC”.

Table 1. Corn Plant Dataset

Transaction ID	Items bought
T1	pen, bread, butter
T2	bread, butter, egg, milk
T3	spinach, egg, milk
T4	bread, butter
T5	bread, butter, ketchup, egg, milk

Trace the results of using the Apriori algorithm on the data above with support threshold s=28% and confidence threshold c=70%.

a. Show the candidate and frequent itemsets for each database scan. Enumerate all the final frequent itemsets.
b. Also indicate the association rules that are generated and highlight the strong ones, sort them by confidence.

Abdul Azzam Ajhari
2201921962

a. Show the candidate and frequent itemsets for each database scan. Enumerate all the final frequent itemsets.

Count support from itemset consisting of 1 item (C_1) and |T| = Transaction ID

Itemset	Sup. count	Sup. calculation	Support
{pen}	1	$s = \frac{\sigma\{pen\}}{\|T\|} = \frac{1}{5} = 0.2$	20%
{bread}	4	$s = \frac{\sigma\{bread\}}{\|T\|} = \frac{4}{5} = 0.8$	80%
{butter}	4	$s = \frac{\sigma\{butter\}}{\|T\|} = \frac{4}{5} = 0.8$	80%
{egg}	3	$s = \frac{\sigma\{egg\}}{\|T\|} = \frac{3}{5} = 0.6$	60%
{milk}	3	$s = \frac{\sigma\{milk\}}{\|T\|} = \frac{3}{5} = 0.6$	60%
{spinach}	1	$s = \frac{\sigma\{spinach\}}{\|T\|} = \frac{1}{5} = 0.2$	20%
{ketchup}	1	$s = \frac{\sigma\{ketchup\}}{\|T\|} = \frac{1}{5} = 0.2$	20%

Find itemset that meets the minimum support requirements s = 28%. The itemset collection is named L_1.

C_1

Itemset	Sup. count	Support
{pen}	1	20%
{bread}	4	80%
{butter}	4	80%
{egg}	3	60%
{milk}	3	60%
{spinach}	1	20%
{ketchup}	1	20%

L_1

Itemset	Sup. count	Support
{bread}	4	80%
{butter}	4	80%
{egg}	3	60%
{milk}	3	60%

Abdul Azzam Ajhari
2201921962

Count support from two itemset consisting of 2 items (C_2) and |T| = Transaction ID

Itemset	Sup. count	Sup. calculation	Support
{bread, butter}	4	$s = \frac{\sigma\{bread,butter\}}{\|T\|} = \frac{4}{5} = 0.8$	80%
{bread, egg}	2	$s = \frac{\sigma\{bread,egg\}}{\|T\|} = \frac{2}{5} = 0.4$	40%
{bread, milk}	2	$s = \frac{\sigma\{bread,milk\}}{\|T\|} = \frac{2}{5} = 0.4$	40%
{butter, egg}	2	$s = \frac{\sigma\{butter,egg\}}{\|T\|} = \frac{2}{5} = 0.4$	40%
{butter, milk}	2	$s = \frac{\sigma\{butter,milk\}}{\|T\|} = \frac{2}{5} = 0.4$	40%
{egg, milk}	3	$s = \frac{\sigma\{egg,milk\}}{\|T\|} = \frac{3}{5} = 0.6$	60%

Find itemset that meets the minimum support requirements s = 28%. The itemset collection is named L_2.

C_2

Itemset	Sup. count	Support
{bread, butter}	4	80%
{bread, egg}	2	40%
{bread, milk}	2	40%
{butter, egg}	2	40%
{butter, milk}	2	40%
{egg, milk}	3	60%

L_2

Itemset	Sup. count	Support
{bread, butter}	4	80%
{bread, egg}	2	40%
{bread, milk}	2	40%
{butter, egg}	2	40%
{butter, milk}	2	40%
{egg, milk}	3	60%

Count support from three itemset consisting of 3 items (C_3) and |T| = Transaction ID

Itemset	Sup. count	Sup. calculation	Support
{bread, butter, egg}	2	$s = \frac{\sigma\{bread,butter,egg\}}{\|T\|} = \frac{2}{5} = 0.4$	40%
{bread, butter, milk}	2	$s = \frac{\sigma\{bread,butter,milk\}}{\|T\|} = \frac{2}{5} = 0.4$	40%
{bread, egg, milk}	2	$s = \frac{\sigma\{bread,egg,milk\}}{\|T\|} = \frac{2}{5} = 0.4$	40%
{butter, egg, milk}	2	$s = \frac{\sigma\{butter,egg,milk\}}{\|T\|} = \frac{2}{5} = 0.4$	40%

Find itemset that meets the minimum support requirements s = 28%. The itemset collection is named L_3.

C_3

Itemset	Sup. count	Support
{bread, butter, egg}	2	40%
{bread, butter, milk}	2	40%
{bread, egg, milk}	2	40%
{butter, egg, milk}	2	40%

L_3

Itemset	Sup. count	Support
{bread, butter, egg}	2	40%
{bread, butter, milk}	2	40%
{bread, egg, milk}	2	40%
{butter, egg, milk}	2	40%

Count support from three itemset consisting of 4 items (C_4) and |T| = Transaction ID

Itemset	Sup. count	Sup. calculation	Support
{bread, butter, egg, milk}	2	$s = \frac{\sigma\{bread,butter,egg,milk\}}{\|T\|} = \frac{2}{5} = 0.4$	40%

Find itemset that meets the minimum support requirements s = 28%. The itemset collection is named L_4.

C_4

Itemset	Sup. count	Support
{bread, butter, egg, milk}	2	40%

L_4

Itemset	Sup. count	Support
{bread, butter, egg, milk}	2	40%

Abdul Azzam Ajhari
2201921962

Final frequent itemsets that meets the minimum requirement of support s = 28%:

{bread, butter, egg, milk} = 40%

{bread, butter, egg} = 40%
{bread, butter, milk} = 40%
{bread, egg, milk} = 40%
{butter, egg, milk} = 40%

{bread, butter} = 80%
{bread, egg} = 40%
{bread, milk} = 40%
{butter, egg} = 40%
{butter, milk} = 40%
{egg, milk} = 60%

{bread} = 80%
{butter} = 80%
{egg} = 60%
{milk} = 60%

b. Also indicate the association rules that are generated and highlight the strong ones, sort them by confidence.

Count confidence from two itemsets and determine if it has a value that exceeds the confidence threshold c = 70% as a TRUE and if not as a FALSE.

Itemset	LEFT	RIGHT	Sup. Count Itemset	Sup. Count LEFT	Calc. Confidence	Confidence	Status
{bread, butter}	bread	butter	4	4	4/4 = 1	100%	TRUE
{bread, butter}	butter	bread	4	4	4/4 = 1	100%	TRUE
{bread, egg}	bread	egg	2	4	2/4 = 0.5	50%	FALSE
{bread, egg}	egg	bread	2	3	2/3 = 0.67	67%	FALSE
{bread, milk}	bread	milk	2	4	2/4 = 0.5	50%	FALSE
{bread, milk}	milk	bread	2	3	2/3 = 0.67	67%	FALSE
{butter, egg}	butter	egg	2	4	2/4 = 0.5	50%	FALSE
{butter, egg}	egg	butter	2	3	2/3 = 0.67	67%	FALSE
{butter, milk}	butter	milk	2	4	2/4 = 0.5	50%	FALSE
{butter, milk}	milk	butter	2	3	2/3 = 0.67	67%	FALSE
{egg, milk}	egg	milk	3	3	3/3 = 1	100%	TRUE
{egg, milk}	milk	egg	3	3	3/3 = 1	100%	TRUE

Abdul Azzam Ajhari
2201921962

Count confidence from three itemsets and determine if it has a value that exceeds the confidence threshold c = 70% as a TRUE and if not as a FALSE.

Itemset	LEFT	RIGHT	Sup. Count Itemset	Sup. Count LEFT	Calc. Confidence	Confidence	Status
{bread, butter, egg}	bread, butter	egg	2	4	2/4 = 0.5	50%	FALSE
{bread, butter, egg}	egg	bread, butter	2	3	2/3 = 0.67	67%	FALSE
{bread, butter, egg}	bread, egg	butter	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg}	butter	bread, egg	2	4	2/4 = 0.5	50%	FALSE
{bread, butter, egg}	butter, egg	bread	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg}	bread	butter, egg	2	4	2/4 = 0.5	50%	FALSE
{bread, butter, milk}	bread, butter	milk	2	4	2/4 = 0.5	50%	FALSE
{bread, butter, milk}	milk	bread, butter	2	3	2/3 = 0.67	67%	FALSE
{bread, butter, milk}	bread, milk	butter	2	2	2/2 = 1	100%	TRUE
{bread, butter, milk}	butter	bread, milk	2	4	2/4 = 0.5	50%	FALSE
{bread, butter, milk}	butter, milk	bread	2	2	2/2 = 1	100%	TRUE
{bread, butter, milk}	bread	butter, milk	2	4	2/4 = 0.5	50%	FALSE
{bread, egg, milk}	bread, egg	milk	2	2	2/2 = 1	100%	TRUE
{bread, egg, milk}	milk	bread, egg	2	3	2/3 = 0.67	67%	FALSE
{bread, egg, milk}	bread, milk	egg	2	2	2/2 = 1	100%	TRUE
{bread, egg, milk}	egg	bread, milk	2	3	2/3 = 0.67	67%	FALSE
{bread, egg, milk}	egg, milk	bread	2	3	2/3 = 0.67	67%	FALSE
{bread, egg, milk}	bread	egg, milk	2	4	2/4 = 0.5	50%	FALSE
{butter, egg, milk}	butter, egg	milk	2	2	2/2 = 1	100%	TRUE
{butter, egg, milk}	milk	butter, egg	2	3	2/3 = 0.67	67%	FALSE
{butter, egg, milk}	butter, milk	egg	2	2	2/2 = 1	100%	TRUE
{butter, egg, milk}	egg	butter, milk	2	3	2/3 = 0.67	67%	FALSE
{butter, egg, milk}	egg, milk	butter	2	3	2/3 = 0.67	67%	FALSE
{butter, egg, milk}	butter	egg, milk	2	4	2/4 = 0.5	50%	FALSE

Abdul Azzam Ajhari
2201921962

Count confidence from four itemsets and determine if it has a value that exceeds the confidence threshold c = 70%.

Itemset	LEFT	RIGHT	Sup. Count Itemset	Sup. Count LEFT	Calc. Confidence	Confidence	Status
{bread, butter, egg, milk}	bread, butter, egg	milk	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg, milk}	bread, butter, milk	egg	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg, milk}	bread, egg, milk	butter	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg, milk}	butter, egg, milk	bread	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg, milk}	milk	bread, butter, egg	2	3	2/3 = 0.67	67%	FALSE
{bread, butter, egg, milk}	egg	bread, butter, milk	2	3	2/3 = 0.67	67%	FALSE
{bread, butter, egg, milk}	butter	bread, egg, milk	2	4	2/4 = 0.5	50%	FALSE
{bread, butter, egg, milk}	bread	butter, egg, milk	2	4	2/4 = 0.5	50%	FALSE
{bread, butter, egg, milk}	bread, butter	egg, milk	2	4	2/4 = 0.5	50%	FALSE
{bread, butter, egg, milk}	egg, milk	bread, butter	2	3	2/3 = 0.67	67%	FALSE
{bread, butter, egg, milk}	bread, egg	butter, milk	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg, milk}	butter, milk	bread, egg	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg, milk}	bread, milk	butter, egg	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg, milk}	butter, egg	bread, milk	2	2	2/2 = 1	100%	TRUE

Abdul Azzam Ajhari
2201921962

Short rules by confidence more than threshold > 70%

Itemset	LEFT	RIGHT	Sup. Count Itemset	Sup. Count LEFT	Calc. Confidence	Confidence	Status
{bread, butter, egg, milk}	bread, butter, egg	milk	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg, milk}	bread, butter, milk	egg	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg, milk}	bread, egg, milk	butter	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg, milk}	butter, egg, milk	bread	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg, milk}	bread, egg	butter, milk	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg, milk}	butter, milk	bread, egg	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg, milk}	bread, milk	butter, egg	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg, milk}	butter, egg	bread, milk	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg}	bread, egg	butter	2	2	2/2 = 1	100%	TRUE
{bread, butter, egg}	butter, egg	bread	2	2	2/2 = 1	100%	TRUE
{bread, butter, milk}	bread, milk	butter	2	2	2/2 = 1	100%	TRUE
{bread, butter, milk}	butter, milk	bread	2	2	2/2 = 1	100%	TRUE
{bread, egg, milk}	bread, egg	milk	2	2	2/2 = 1	100%	TRUE
{bread, egg, milk}	bread, milk	egg	2	2	2/2 = 1	100%	TRUE
{butter, egg, milk}	butter, egg	milk	2	2	2/2 = 1	100%	TRUE
{butter, egg, milk}	butter, milk	egg	2	2	2/2 = 1	100%	TRUE
{bread, butter}	bread	butter	4	4	4/4 = 1	100%	TRUE
{bread, butter}	butter	bread	4	4	4/4 = 1	100%	TRUE
{egg, milk}	egg	milk	3	3	3/3 = 1	100%	TRUE
{egg, milk}	milk	egg	3	3	3/3 = 1	100%	TRUE

There's 20 rules above the confident treshold