Applying Machine Learning Techniques to Retail Data

Janani Ravi
Co-founder, Loonycorn

www.loonycorn.com

Overview

Association rules learning

Frequent itemsets and support

Evaluating association rules using confidence, lift, and conviction

Market basket analysis using the Apriori algorithm

Association Rules Learning

Association Rules Learning

Data mining technique usually used to identify interesting patterns in which items appear together - for instance beer and diapers in shopping baskets.

Rules and Strong Rules

Rules are of the form “If X then Y”

Strong rules are rules supported by probability

Strong rules can be extremely useful

- Recommendations
- Cross-sell
- Up-sell

Market Basket Analysis

Classic use for association rules learning

Used to identify items sold together

- People who buy diapers also buy beer

Also used to segment users

- People who like diapers but not beer

Related to recommendation systems

Evaluating Association Rules

Association Rules Learning

Data mining technique usually used to identify interesting patterns in which items appear together.

Association Rules Learning

Data mining technique usually used to identify strong rules in data.

Strong Rules

Rules with support and confidence that exceed thresholds for minimum support and minimum confidence respectively.

Strong Rules

Rules with support and confidence that exceed thresholds for **minimum support** and **minimum confidence** respectively.

Evaluating Association Rules

Support

Confidence

Lift

Conviction

Evaluating Association Rules

Support	**Confidence**
Lift	**Conviction**

Support is a metric defined for itemsets and not association rules

Supermarket Transactions

Each row represents one market basket at check-out

#	Milk	Bread	Butter	Beer	Diapers
1	1	1	0	0	0
2	0	0	1	0	0
3	0	0	0	1	1
4	1	1	1	0	0
5	0	1	0	0	0

{Milk, Bread, Butter, Beer, Diapers}

Items

$I = \{I_1, I_2, I_3, I_4, I_5\}$

{Milk, Bread}

{Bread, Butter}

{Beer, Diapers}

Different Itemsets

Each itemset contains some subset of the master set of all items

Supermarket Transactions

Each row represents one market basket at check-out

#	Milk	Bread	Butter	Beer	Diapers
1	1	1	0	0	0
2	0	0	1	0	0
3	0	0	0	1	1
4	1	1	1	0	0
5	0	1	0	0	0

T_1 = {Milk, Bread}

T_2 = {Butter}

T_3 = {Beer, Diapers}

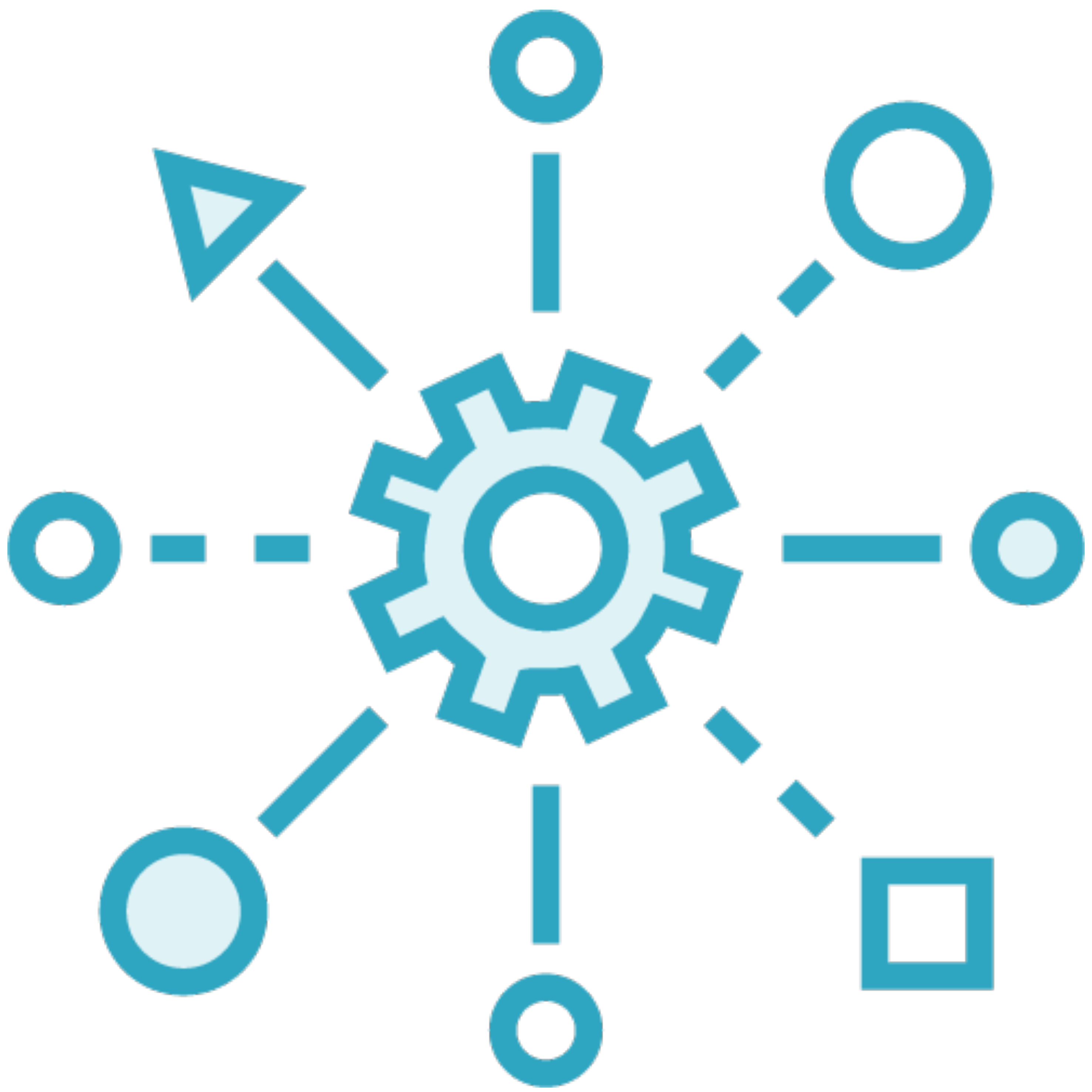
T_4 = {Milk, Bread, Butter}

T_5 = {Bread}

Itemset representation

Each transaction represents an itemset

Frequent Itemsets



A frequent itemset is one that appears frequently

i.e. exceeds threshold for probability of occurrence

Probability of occurrence of an itemset is called the **Support**

Support

Measure of how often an itemset appears in the data.

Support

Measure of how significant or important an itemset is in the data.

Metrics Associated with Rules

Support	Confidence
Lift	Conviction

{Beer} = > {Diapers}

Rule: A = > C

"If a transaction contains A, it will contain C"

{Beer} = > {Diapers}

Rule: A = > C

There is a directionality associated with rules

{Diaper} = > {Beers}

Rule: C = > A

Metric for a rule A->C is not the same as the metric of a rule C->A

A: Antecedents
C: Consequents

Evaluating Association Rules

Support

Confidence

Lift

Conviction

Confidence

Confidence of a rule A = > C, where A and C are itemsets, is the proportion of transactions containing A that also contain C

Confidence

Confidence of a rule **A = > C**, where A and C are itemsets, is the proportion of transactions containing A that also contain C

Confidence

Confidence of a rule A = > C, is the probability of seeing the itemset C in a transaction where the transaction also contains A

Confidence

Confidence of a rule A = > C, is the **probability of seeing the itemset C in a transaction where the transaction also contains A**

Evaluating Association Rules

Support

Confidence

Lift

Conviction

Lift

Lift of a rule A = > C, where A and C are itemsets, is the increase in probability of C if A has occurred vs. probability of C if we know nothing about A

Lift

Lift of a rule A = > C, where A and C are itemsets, **is the increase in probability of C if A has occurred** vs. probability of C if we know nothing about A

Lift

Lift of a rule A = > C, where A and C are itemsets, is the increase in probability of C if A has occurred vs. **probability of C if we know nothing about A**

Lift

How much more often do A and C occur together than we would expect if they were statistically independent.

Evaluating Association Rules

Support	Confidence
Lift	**Conviction**

Conviction

Conviction of a rule A = > C, measures, of the occasions when the rule is false (A and C do not co-occur), how often C does not occur at all.

Conviction

Conviction of a rule A = > C, measures, **of the occasions when the rule is false (A and C do not co-occur)**, how often C does not occur at all.

Conviction

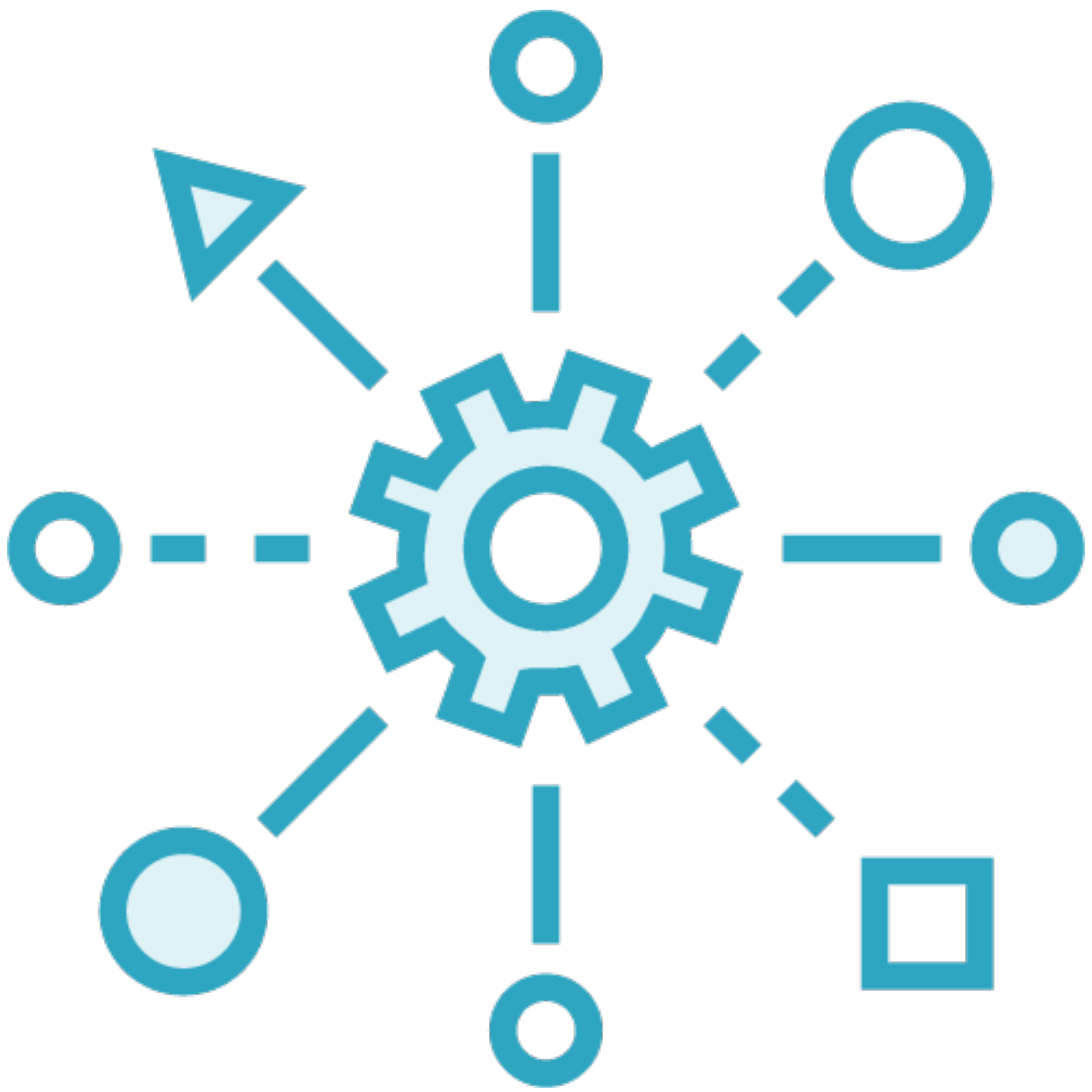
Conviction of a rule A = > C, measures, of the occasions when the rule is false (A and C do not co-occur), **how often C does not occur at all.**

Conviction

A measure of how dependent C is on A.

Apriori Algorithm

Frequent Itemsets



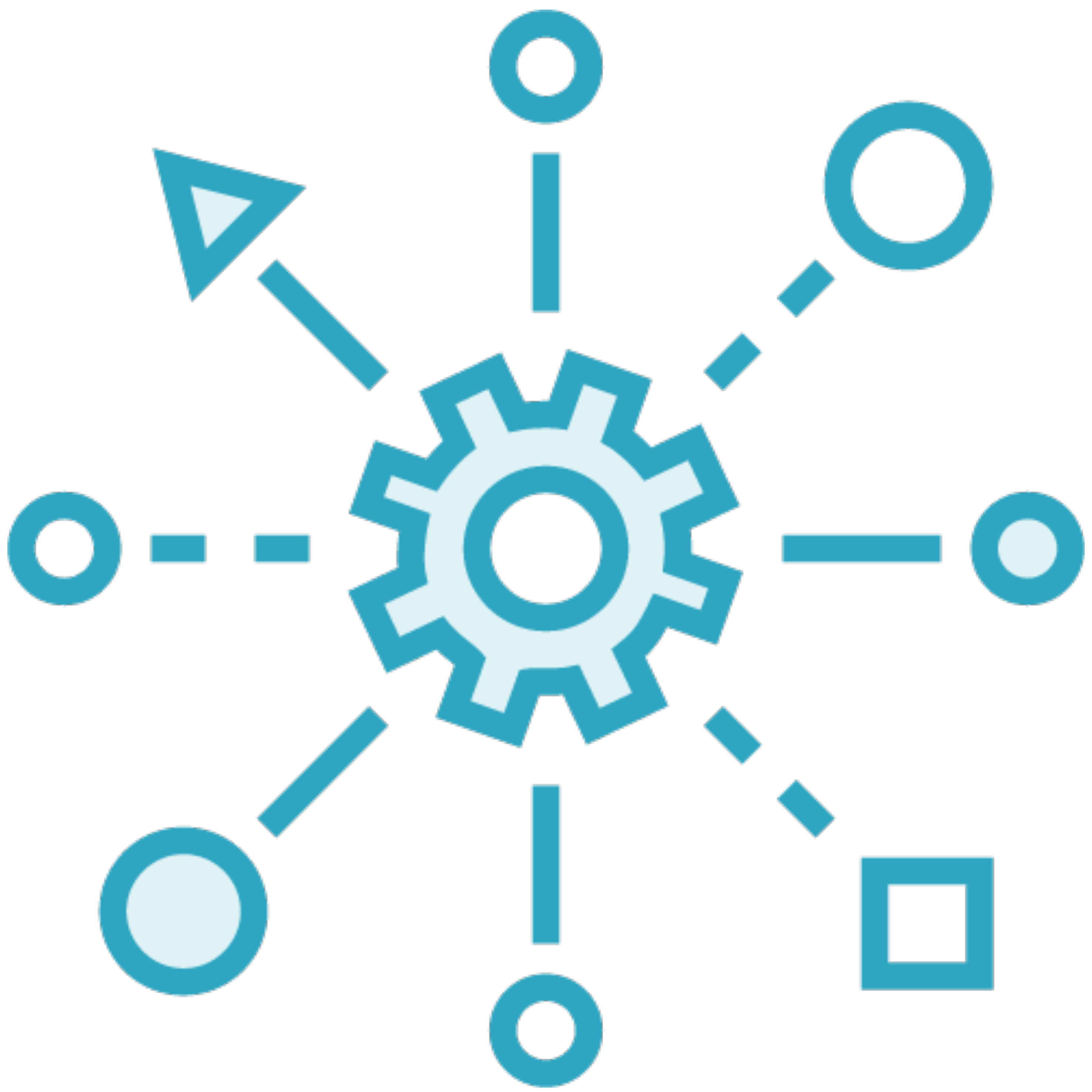
Finding frequency of an itemset can involve searching set of all subsets

Known as powerset

For set with N elements, powerset has 2^N-1 elements

Not computationally scalable

Frequent Itemsets

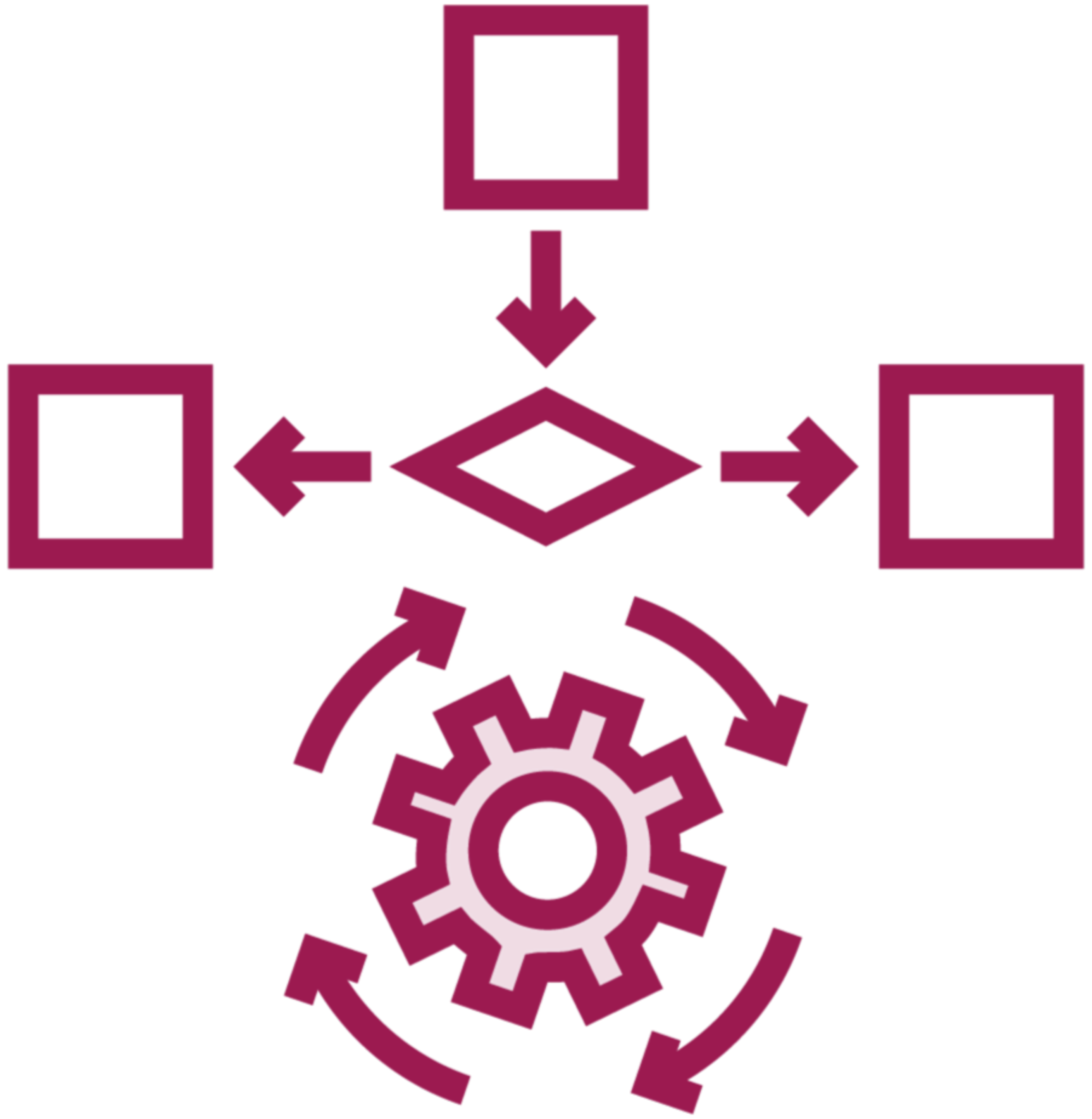


Downward-closure property of itemsets

All subsets of a frequent itemset are also frequent itemsets

Allows efficient computation of frequent itemsets

Apriori Algorithm



- Uses breadth-first search and hash tree for efficient implementation
- Bottom-up approach to generating candidate frequent sets
- Starts with 1-item itemsets
- Keeps extending them if possible
- Stops when no longer possible

Demo

Performing market basket analysis using the Apriori algorithm and association rules learning

Summary

Association rules learning

Frequent itemsets and support

Evaluating association rules using confidence, lift, and conviction

Market basket analysis using the Apriori algorithm

Resources Referenced in This Course

Gartner reports

https://www.gartner.com/doc/reprints?id=1-265MCE7K&ct=210528&st=sb
https://www.gartner.com/en/newsroom/press-releases/2020-02-04-gartner-predicts-at-least-two-top-global-retailers-wi

Use cases for ML in retail:

https://spd.group/artificial-intelligence/ai-for-retail/#AI_in_the_Retail_Supply_Chain
https://tryolabs.com/guides/retail-innovations-machine-learning

Challenges for ML in retail:

https://www.jbssolutions.com/resources/blog/challenges-and-pitfalls-retailers-face-starting-machine-learning/

Resources Referenced in This Course

Price elasticity of demand
https://hbr.org/2015/08/a-refresher-on-price-elasticity

Case Study: Price Optimization in Fashion E-commerce
https://arxiv.org/pdf/2007.05216v2.pdf

Case Study: Dynamic Vehicle Routing Problem
https://arxiv.org/pdf/2008.11719.pdf

Related Courses

Machine Learning for Financial Services

Machine Learning for Healthcare