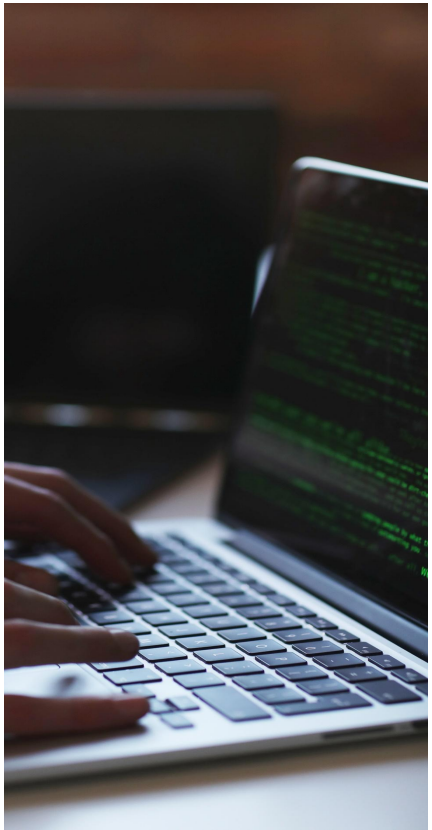
Course-Net
Empowering IT People

Introduction to Association Rules

Course-Net
Empowering IT People

EmpoweringITPeople

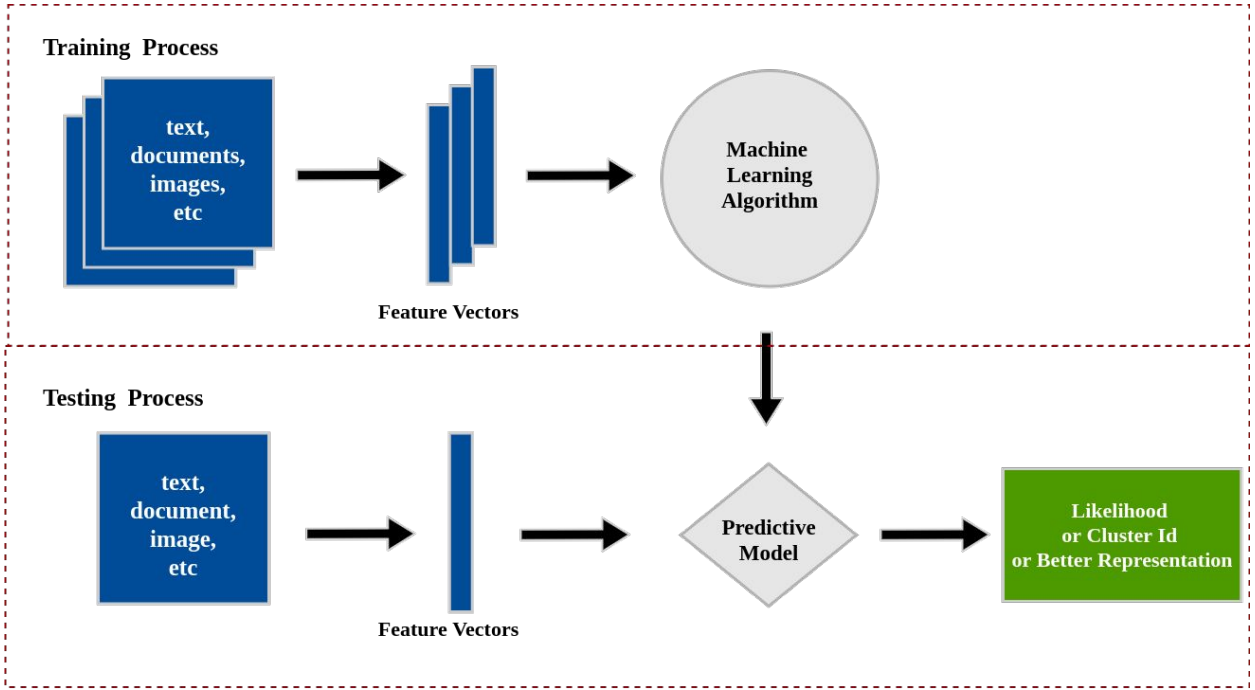


APA YANG AKAN KITA PELAJARI

Associations and applications

- What is Association Rules
- Association Rules Intuition
- Association Rules Algorithm
- Applications of Association
- Association Rules Cases

Unsupervised Learning Process



Course-Net
Empowering IT People

#EmpoweringITPeople

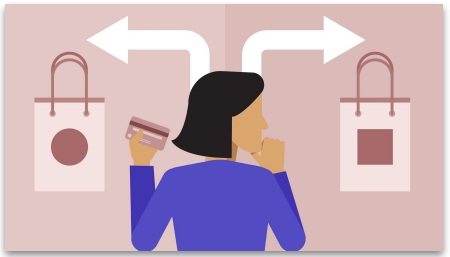
Apa itu Association Rules Mengapa Kita Pelajari di Data Science?

Association Rules Learning: The unsupervised learning technique that checks for the dependency of one data item on another data item and maps accordingly so that it can be more

Association Rules Use Cases

Market Basket Analysis

Recommendation Systems



Customer Behavior

Market Basket Transaction

Purchase ID	Items
1	Bread, Milk
2	Bread, Diaper, Beer, Snacks
3	Milk, Diaper, Beer, Soda
4	Bread, Milk, Diaper, Beer
5	Bread, Milk, Diaper, Soda

Association Rules:

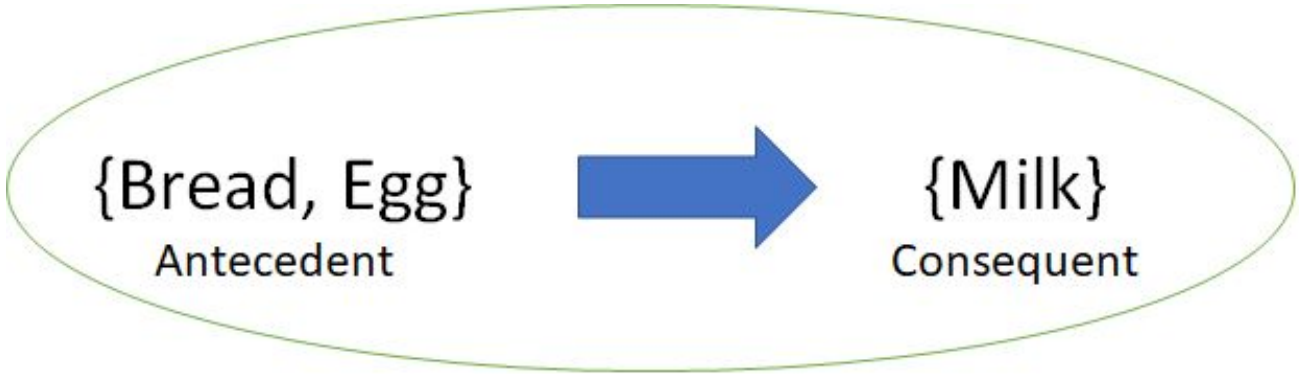
{Bread} -> {Snacks}

{Bread, Beer} -> {Snacks}

{Bread, Milk} -> {Diaper, Soda}

{Milk, Bread} -> {Soda, Beer}

Itemset



Itemset = {Bread, Egg, Milk}

Association Rules Evaluation

Metrics for Evaluation:

1. Support
2. Confidence

Support

The relative frequency that the rules show up

$$Support(\{X\} \rightarrow \{Y\}) = \frac{Transactions\ containing\ both\ X\ and\ Y}{Total\ number\ of\ transactions}$$

Purchase ID	Items
1	Bread, Milk
2	Bread, Diaper, Beer, Snacks
3	Milk, Diaper, Beer, Soda
4	Bread, Milk, Diaper, Beer
5	Bread, Milk, Diaper, Soda

{Bread, Diaper} => Beer

S({Bread, Diaper, Beer}) = 2/5

Support Count = 2

Measure of the reliability of a rule by likeliness of occurrence of consequent

$$Confidence(\{X\} \rightarrow \{Y\}) = \frac{Transactions\ containing\ both\ X\ and\ Y}{Transactions\ containing\ X}$$

Purchase ID	Items
1	Bread, Milk
2	Bread, Diaper, Beer, Snacks
3	Milk, Diaper, Beer, Soda
4	Bread, Milk, Diaper, Beer
5	Bread, Milk, Diaper, Soda

{Bread, Diaper} => Beer

C({Bread, Diaper, Beer}) = 2/3

Association Rules Mining

From given the dataset, find the rules that:

1. Support >= min support threshold
2. Confidence >= min confidence threshold

Mining Association Rules

1. **Frequent Itemset Generation**
   - Generate all itemsets whose support >= min support

1. **Rule Generation**
   - Generate high confidence rules from each frequent itemset

Association Rules Algorithm

1. Apriori
2. Eclat

Apriori Steps

1. Set minimum Support and Confidence threshold
   Repeat until there are no new frequent itemsets:
1. Takes all subset in the loop in the transactions which have bigger Support than minimum Confidence
2. Take all the rules of these subsets which have bigger confidence than minimum confidence
3. Sort the rules by decreasing lift

Lift = (Confidence(X-> Y) / Support(Y))