Association Rule

Created Date: 2022-11-18

> Metadata 🧰
>
> - Title: Association Rule
> - Author: Andrew Jones
> - Reference: Data Science Infinity

> Links & Tags 🖇️
>
> - Index: Course Note Index
> - Atomic Tag: #datascience
> - Subatomic Tags: #machinelearning #associationrule #unsupervisedlearning

---

High-Level Overview

Jupyter Notebook: Advanced Association Rule Template

> Association Rule Learning is an approach that discovers the strength of relationships between different data points. It is commonly used to understand which product are frequently purchased together.

- A commonly used algorithm to measure the strength of connections is Apriori
- Apriori contains four key metrics;
  - Support: % of all transactions that contain both item A & item B
  - Confidence: Of all transactions that included item A, what proportion also included item B

    - Expected Confidence: % of all transactions that contain item B
    - Lift: The factor by which Confidence exceeds the Expected Confidence (controls of item popularity)
    - *Item relationships that have a high Lift score, but low Support score should be interpreted with caution*
- We can use a matrix to visualize these metrics
- Uses of Association Rule Learning;
    - How to preset or arrange items in the store
    - Understand if discounts should be offered on items
    - Understand if only one the items should be advertised at one time

Formulas

- $Support = \frac{Num\ Transactions\ Containing\ Items\ A\ and\ B}{Total\ Transactions}$
- $Confidence = \frac{Num\ Transactions\ Containing\ Items\ A\ and\ B}{Num\ Transactions\ Containing\ Item\ A}$
- $Expected\ Confidence = Support = \frac{Num\ Transactions\ Containing\ Item\ B}{Total\ Transactions}$
- $Lift = \frac{Confidence}{Expected\ Confidence}$