

PRODUCT RECOMMENDATION

Cross-selling models & Recommendation systems

Timpika Aungnoi 6310412025

MARKET BASKET ANALYSIS



is **frequent itemset** mining. This process identifies customer buying habits by finding associations between the different items that customers place in their “shopping baskets”

For example, if customers are buying milk, how probably are they to also buy bread (and which kind of bread) on the same trip to the supermarket?

Analysis of buying patterns that reflect items that are frequently associated or bought together. Such patterns will be represented in the form of **association rules**.

References:

<https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-on-market-basket-analysis/>

ASSOCIATION RULE

Finding frequent patterns, associations, correlations, or causal structures among sets of items in transaction databases.

The rule could be written as

If {A} Then {B}

The If part of the rule (the {A} above) is known as the antecedent and the THEN part of the rule is known as the consequent (the {B} above). The antecedent is the condition and the consequent is the result.

References:

https://paginas.fe.up.pt/~ec/files_1112/week_04_Association.pdf

<https://infocenter.informationbuilders.com/wf80/index.jsp?topic=%2Fpubdocs%2FRStat16%2Fsource%2Ftopic49.htm>

THREE MEASURES OF THE ASSOCIATION RULE

- 1. SUPPORT** IS THE NUMBER OF TRANSACTIONS THAT INCLUDE ITEMS IN THE {A} AND {B} PARTS OF THE RULE AS A PERCENTAGE OF THE TOTAL NUMBER OF TRANSACTIONS. IT IS A MEASURE OF HOW FREQUENTLY THE COLLECTION OF ITEMS OCCURS TOGETHER AS A PERCENTAGE OF ALL TRANSACTIONS.
- 2. CONFIDENCE** IS THE NUMBER OF TRANSACTIONS THAT INCLUDE ITEMS IN THE {A} AND {B} PARTS OF THE RULE AS A PERCENTAGE OF THE TOTAL NUMBER OF TRANSACTIONS. IT IS A MEASURE OF HOW FREQUENTLY THE COLLECTION OF ITEMS OCCURS TOGETHER AS A PERCENTAGE OF ALL TRANSACTIONS.
- 3. LIFT** IS THE RATIO OF CONFIDENCE TO EXPECTED CONFIDENCE. EXPECTED CONFIDENCE IS THE CONFIDENCE DIVIDED BY THE FREQUENCY OF B. THE LIFT TELLS US HOW MUCH BETTER A RULE IS AT PREDICTING THE RESULT THAN JUST ASSUMING THE RESULT IN THE FIRST PLACE. GREATER LIFT VALUES INDICATE STRONGER ASSOCIATIONS.

References:

<https://infocenter.informationbuilders.com/wf80/index.jsp?topic=%2Fpubdocs%2FRStat16%2Fsource%2Ftopic49.htm>

DATASET

SURVEY QUESTIONNAIRE IN CLASS ABOUT PRODUCTS THAT STUDENTS BOUGHT. (N=29)

PROD_REC.CSV

OBJECTIVE

USING MARKET BASKET ANALYSIS TO DETERMINE WHAT ITEMS ARE PURCHASED TOGETHER,
PURCHASED SEQUENTIALLY.



CONCLUSION

EXAMPLE OF PRODUCT RECOMMENDATIONS FROM RESULTS

{RUNNING SHOES} -> {BLUETOOTH HEADPHONES, AIRPODS, TRUE WIRELESS}

IT MEANS IF THEY BUY RUNNING SHOES, THEY MORE LIKELY TO BUY BLUETOOTH HEADPHONES, AIRPODS, AND TRUE WIRELESS

SUPPORT = 0.76 WHICH MEANS 76% OF CUSTOMERS PURCHASED RUNNING SHOES AND BLUETOOTH HEADPHONES, AIRPODS, AND TRUE WIRELESS

CONFIDENCE = 0.79 WHICH MEANS 79% OF THE CUSTOMERS THAT BOUGHT RUNNING SHOES ALSO BOUGHT BLUETOOTH HEADPHONES, AIRPODS, AND TRUE WIRELESS

LIFT = 1.04 WHICH MEANS THE 4% INCREASE IN THE EXPECTATION THAT SOMEONE WILL BUY BLUETOOTH HEADPHONES, AIRPODS, AND TRUE WIRELESS, WHEN WE KNOW THAT THEY BOUGHT RUNNING SHOES.

RESULT FROM PYTHON

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(รองเท้าวิ่ง)	(หูฟัง bluetooth (airpods,truewireless))	0.965517	0.758621	0.758621	0.785714	1.035714	0.026159	1.126437
1	(หูฟัง bluetooth (airpods,truewireless))	(รองเท้าวิ่ง)	0.758621	0.965517	0.758621	1.000000	1.035714	0.026159	inf
2	(รองเท้าวิ่ง)	(Ebook)	0.965517	0.655172	0.655172	0.678571	1.035714	0.022592	1.072797
3	(Ebook)	(รองเท้าวิ่ง)	0.655172	0.965517	0.655172	1.000000	1.035714	0.022592	inf
4	(เบาะรองนั่ง)	(คอร์สเรียนเสริม)	0.551724	0.931034	0.551724	1.000000	1.074074	0.038050	inf
...
103	(คอร์สเรียนเสริม, Ebook)	(รองเท้าวิ่ง, เคส ipad)	0.620690	0.862069	0.586207	0.944444	1.095556	0.051130	2.482759
104	(รองเท้าวิ่ง)	(เคส ipad, คอร์สเรียนเสริม, Ebook)	0.965517	0.586207	0.586207	0.607143	1.035714	0.020214	1.053292
105	(เคส ipad)	(รองเท้าวิ่ง, คอร์สเรียนเสริม, Ebook)	0.896552	0.620690	0.586207	0.653846	1.053419	0.029727	1.095785
106	(คอร์สเรียนเสริม)	(รองเท้าวิ่ง, เคส ipad, Ebook)	0.931034	0.620690	0.586207	0.629630	1.014403	0.008323	1.024138
107	(Ebook)	(รองเท้าวิ่ง, เคส ipad, คอร์สเรียนเสริม)	0.655172	0.793103	0.586207	0.894737	1.128146	0.066587	1.965517

RESULT FROM PYTHON

