빅데이터와 비즈니스 모델 1차 과제

**A case on Association Rule Analysis**

20191245 노유정

The Association Rules is method that analyze how related the two events A and C are. It calculates association, not causal relationship, as it does not consider the time difference. There are two important concept in association rule. One is the confidence(Pr(C|A) and the other is the Lift Ratio(Pr(C|A) / Pr(C). The higher the Lift(Confidence), the better the rule.

It is difficult to randomly search for the transaction file(data file that records the selected item in detail) of the measuring full rule . The method used at this time is A priori Algorithm. The general procedure for setting the priori algorithm is as follows. The preliminary algorithm begins with determining support (the number of transactions of specific kits appearing in the transaction file) of 1-item sets (set of 1 item of interest). If any 1-item set does not meet the given criteria, the item is dismissed. The next step is to create a 2-item set (with the smallest support) with a 1-item set that has not been rejected. The 2-item set that does not have minimal support is dismissed. At this time, the 2-item set belongs to the concept, and the other belongs to the antecedent.

There are two tuning parameters in designing the association rules. When one side is determined, the other is determined automatically. However, the rule, which is the setting of minimal support and confidence, is different from other settings, so you have to try several times to get a meaningful result. Of course, even if the priorio algorithm creates a strong association rule, it cannot be adopted if the rule is not interpretable.