**Explanation of Support, Confidence, and Lift in Association Rule Analysis**

연관 규칙 분석에서 사용되는 주요 메트릭인 지지도(Support), 신뢰도(Confidence), 향상도(Lift)는 각각 연관 규칙의 유의미성과 강도를 평가하는 데 중요한 역할을 합니다. 이 메트릭들은 다음과 같은 의미를 가지고 있습니다:  
  
1. 지지도 (Support)  
 - 정의: 지지도는 특정 항목 집합(Itemset)이 데이터셋에서 얼마나 자주 나타나는지를 나타냅니다.  
 - 공식:   
 Support(X) = (X를 포함하는 거래의 수) / (전체 거래 수)  
 - 의미:   
 - 지지도가 높을수록 해당 항목 집합이 전체 거래에서 자주 나타난다는 것을 의미합니다.  
 - 예를 들어, 전체 100건의 거래 중 20건에 "우유"가 포함되어 있다면, 우유의 지지도는 0.2입니다.  
 - 지지도는 규칙의 빈도를 평가하는 기본적인 지표로 사용됩니다.  
  
2. 신뢰도 (Confidence)  
 - 정의: 신뢰도는 연관 규칙에서 선행 조건(Antecedent)이 발생했을 때, 그 결과(Consequent)가 발생할 확률을 나타냅니다.  
 - 공식:  
 Confidence(X → Y) = (X와 Y를 모두 포함하는 거래의 수) / (X를 포함하는 거래의 수)  
 - 의미:  
 - 신뢰도는 선행 조건이 주어졌을 때, 결과가 발생할 가능성을 보여줍니다.  
 - 예를 들어, "우유"를 구매한 거래 중 60%가 "빵"도 함께 구매했다면, "우유 → 빵"의 신뢰도는 0.6입니다.  
 - 신뢰도가 높을수록, 선행 조건이 발생할 때 결과도 자주 발생한다는 것을 의미합니다.  
  
3. 향상도 (Lift)  
 - 정의: 향상도는 두 항목 집합이 독립적일 때에 비해 함께 발생할 가능성이 얼마나 더 높은지를 나타내는 지표입니다.  
 - 공식:  
 Lift(X → Y) = Confidence(X → Y) / Support(Y)  
 - 의미:  
 - Lift > 1: X와 Y가 양의 상관관계를 가집니다. 즉, X가 발생할 때 Y가 발생할 가능성이 평균보다 높습니다. (규칙이 유의미함)  
 - Lift = 1: X와 Y는 독립적입니다. X의 발생이 Y의 발생에 영향을 미치지 않습니다.  
 - Lift < 1: X가 발생할 때 Y가 발생할 가능성이 평균보다 낮습니다. (규칙이 비유의미하거나 반대의 연관성을 가질 수 있음)  
 - 향상도는 규칙의 유의미성을 판단하는 데 중요한 지표로 사용됩니다. 예를 들어, "우유"를 구매한 사람이 "빵"을 구매할 가능성이 전체 거래에서 "빵"이 구매되는 평균 확률보다 2배 높다면, 그 규칙의 Lift는 2입니다.  
  
요약  
- 지지도 (Support): 전체 거래 중 해당 항목 집합이 얼마나 자주 발생하는지를 나타냅니다.  
- 신뢰도 (Confidence): 선행 조건이 발생할 때 결과가 함께 발생할 확률을 나타냅니다.  
- 향상도 (Lift): 두 항목 집합이 독립적일 때에 비해 함께 발생할 가능성이 얼마나 더 높은지를 나타내며, 규칙의 유의미성을 평가하는 데 중요한 지표입니다.  
  
이 세 가지 메트릭을 통해 연관 규칙의 강도와 유의미성을 평가하고, 이를 바탕으로 중요한 인사이트를 도출할 수 있습니다.