

Problem 5: Association Rules – Apriori Algorithm

A database has five transactions. Let $min_sup=60\%$ and $min_confid=80\%$

<i>TID</i>	<i>items_bought</i>
T100	{ D , A , C , E , B }
T200	{ K , A , G , D , B }
T300	{ B , A , D }
T400	{ D , F , B }
T500	{ C , A , B , E , K }

- (a) find all frequent itemsets using *Apriori* algorithm (assume *minimum support* count is **3**)
- (b) list all of the **strong** association rules (with support, s , and confidence, c) matching the following *metarule*, where X is a variable representing customers, and $item_i$ are variables representing items (e.g., “**A**”, “**B**”, etc.).

$$\forall x \in \text{transaction}, buys(X, item_1) \Rightarrow buys(X, item_2) \wedge buys(X, item_3)[s, c]$$
$$\forall x \in \text{transaction}, buys(X, item_1) \wedge buys(X, item_2) \Rightarrow buys(X, item_3)[s, c]$$