Apriori Algorithm

Step 1: Dataset (Transactions)

Suppose we have 5 transactions in a supermarket:

- **T1** = {Milk, Bread, Butter}
- **T2** = {Milk, Bread}
- **T3** = {Milk, Apple}
- **T4** = {Bread, Butter}
- **T5** = {Milk, Bread, Apple}

Step 2: Define Parameters

- **Minimum Support = 2 transactions**
- **Minimum Confidence = 60%**

Step 3: Generate Candidate 1-itemsets (C1)

Count each individual item's frequency:

Item	Count	Support
Milk	4	4/5 = 0.8
Bread	4	0.8
Butter	2	0.4
Apple	2	0.4

✅ All items meet min support (≥2).
So **L1 = {Milk, Bread, Butter, Apple}**

Step 4: Generate Candidate 2-itemsets (C2)

Now we combine L1 items into pairs and count:

Itemset	Count	Support
{Milk, Bread}	3	0.6
{Milk, Butter}	1	0.2
{Milk, Apple}	2	0.4

Itemset	Count	Support
{Bread, Butter}	2	0.4
{Bread, Apple}	1	0.2
{Butter, Apple}	0	0.0

✅ Frequent pairs (≥2): **{Milk, Bread}, {Milk, Apple}, {Bread, Butter}**

So **L2 = {Milk, Bread}, {Milk, Apple}, {Bread, Butter}**

Step 5: Generate Candidate 3-itemsets (C3)

From L2, we can try {Milk, Bread, Apple}, {Milk, Bread, Butter}

- {Milk, Bread, Apple} appears in **T5 only → Count = 1** (not frequent)
- {Milk, Bread, Butter} appears in **T1 only → Count = 1** (not frequent)

❌ No frequent 3-itemsets.

Step 6: Generate Association Rules

Now we make rules from frequent itemsets:

Example: From {Milk, Bread} (Support = 3/5 = 60%)

- Rule: **Milk → Bread**
  - Confidence = Support(Milk ∩ Bread) / Support(Milk)
  - = 3 / 4 = 0.75 (75%) ✅
- Rule: **Bread → Milk**
  - = 3 / 4 = 0.75 (75%) ✅

From {Milk, Apple} (Support = 2/5 = 40%)

- Rule: **Milk → Apple** = 2/4 = 0.5 ❌ (fails confidence)
- Rule: **Apple → Milk** = 2/2 = 1.0 ✅

From {Bread, Butter} (Support = 2/5 = 40%)

- Rule: **Bread → Butter** = 2/4 = 0.5 ❌
- Rule: **Butter → Bread** = 2/2 = 1.0 ✅

✅ **Final Strong Rules**

- **Milk → Bread (75%)**

- **Bread → Milk (75%)**
- **Apple → Milk (100%)**
- **Butter → Bread (100%)**