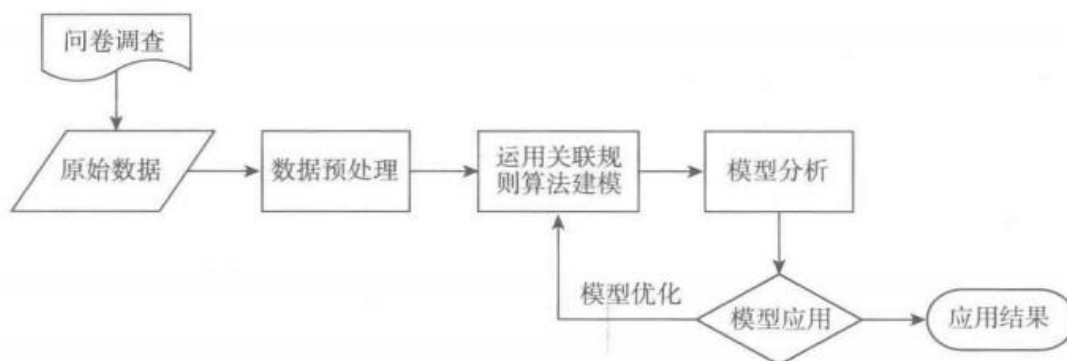


中医证型关联规则挖掘

- 1) 借助三阴乳腺癌患者的病理信息，挖掘患者的症状与中医证型之间的关联关系。
- 2) 对截断治疗提供依据，挖掘潜在性证素。

分析方法与过程



数据获取

数据预处理

数据清洗、属性规约和数据变换。

数据清洗

属性规约

数据变换

属性构造和**数据离散化**对数据进行处理。首先通过属性构造，获得证型系数，然后通过聚类算法对数据进行离散化处理，形成建模数据。

模型构建

由图 8-3 可知，模型主要由输入、算法处理、输出部分组成。输入部分包括：①建模样本数据的输入；②建模参数的输入。算法处理部分是 Apriori 关联规则算法。输出部分为关联规则的结果。

模型具体实现步骤为：首先设置建模参数最小支持度、最小置信度，输入建模样本数据。然后采用 Apriori 关联规则算法对建模的样本数据进行分析，以模型参数设置的最小支持度、最小置信度以及分析目标作为条件，如果所有的规则都不满足条件，则需要重新调整模型参数，否则输出关联规则结果。

目前，如何设置最小支持度与最小置信度，并没有统一的标准。大部分都是根据业务经验设置初始值，然后经过多次调整，获

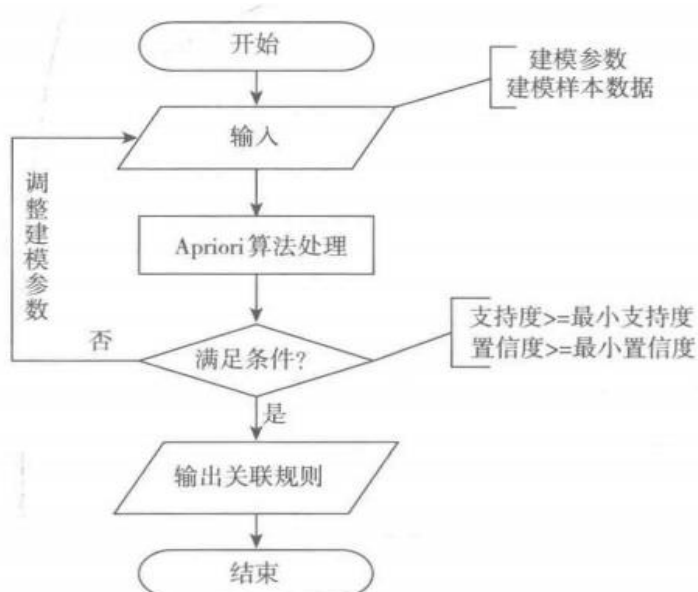


图 8-3 中医证型关联规则模型流程图

取与业务相符的关联规则结果。

模型分析

模型应用