What is Constraint-Based Association Mining?

A data mining procedure can uncover thousands of rules from a given set of information, most of which end up being independent or tedious to the users. Users have a best sense of which "direction" of mining can lead to interesting patterns and the "form" of the patterns or rules they can like to discover.

Therefore, a good heuristic is to have the users defines such intuition or expectations as constraints to constraint the search space. This strategy is called constraint-based mining.

Constraint-based algorithms need constraints to decrease the search area in the frequent itemset generation step (the association rule generating step is exact to that of exhaustive algorithms).

The general constraint is the support minimum threshold. If a constraint is uncontrolled, its inclusion in the mining phase can support significant reduction of the exploration space because of the definition of a boundary inside the search space lattice, following which exploration is not needed.

The important of constraints is well-defined – they create only association rules that are appealing to users. The method is quite trivial and the rules space is decreased whereby remaining methods satisfy the constraints.

Constraint-based clustering discover clusters that satisfy user-defined preferences or constraints. It depends on the characteristics of the constraints, constraint-based clustering can adopt rather than different approaches.

The constraints can include the following which are as follows –

Knowledge type constraints – These define the type of knowledge to be mined, including association or correlation.

Data constraints – These define the set of task-relevant information such as Dimension/level constraints – These defines the desired dimensions (or attributes) of the information, or methods of the concept hierarchies, to be utilized in mining.

Interestingness constraints – These defines thresholds on numerical measures of rule interestingness, including support, confidence, and correlation.

Rule constraints – These defines the form of rules to be mined. Such constraints can be defined as metarules (rule templates), as the maximum or minimum number of predicates that can appear in the rule antecedent or consequent, or as relationships between attributes, attribute values, and/or aggregates.

The following constraints can be described using a high-level declarative data mining query language and user interface. This form of constraint-based mining enables users to define the rules that they can like to uncover, thus by creating the data mining process more efficient.

Furthermore, a sophisticated mining query optimizer can be used to deed the constraints defined by the user, thereby creating the mining process more effective. Constraint-based mining boost interactive exploratory mining and analysis.