A FAST ALGORITHM FOR CONSTRUCTING DECISION TREES

Abstract of the Invention

A method to construct the decision tree using a FAST algorithm is disclosed. The method includes building the prefix tree from the set of records of the dataset. The prefix tree may be traversed to count the frequency of a predictor attribute value stored in each node of the prefix tree. The frequencies counted are stored in a frequency table and the frequency table is used to determine the best split condition of the prefix tree. Upon determination of the best split condition, the prefix tree may be traversed to split the prefix tree at the node of the prefix tree where the best split condition is located. On splitting the prefix tree at least one of a new prefix tree is formed and also the portion of the prefix tree is also retained where the best split condition is not located. The new prefix tree and the retained portion of the prefix tree may be used to construct the decision tree. The construction of the decision tree may be a recursive process until certain set of predefined condition is matched while traversing the prefix tree for splitting the prefix tree.