A Concept of Data Mining with XML Document

Authors : Akshay Agrawal, Anand K. Srivastava

Abstract : The increasing amount of XML datasets available to casual users increases the necessity of investigating techniques to extract knowledge from these data. Data mining is widely applied in the database research area in order to extract frequent correlations of values from both structured and semi-structured datasets. The increasing availability of heterogeneous XML sources has raised a number of issues concerning how to represent and manage these semi structured data. In recent years due to the importance of managing these resources and extracting knowledge from them, lots of methods have been proposed in order to represent and cluster them in different ways.