

Data Mining Methods using SQL server resources

Abstract

SQL Server has been a leader in predictive analytics since the 2000 release, by providing data mining in Analysis Services. SQL Server Data Mining provides an integrated platform for predictive analytics that encompasses data cleansing and preparation, machine learning, and reporting. SQL Server Data Mining includes multiple standard algorithms, including EM and K-means clustering models, neural networks, logistic regression and linear regression, decision trees, and naive bayes classifiers. All models have integrated visualizations to help you develop, refine, and evaluate your models.

Today there are many algorithms and methods used for analysis of data. The main problem is to find a suitable algorithm for data retrieval in case of a particular problem. The work is examined and analyzed by the factors influencing the selection of the algorithm.

Master's thesis may be useful for scientific workers, teachers, doctoral students, graduate students, bachelors and students of higher education institutions. Its main conclusions can be used to conduct further studies in various organizations, as well as in pedagogic practice.

Keywords— SQL Server Data mining, k-means clustering, Decision Tree

Julia Shavkatsishvili