Abstract

Healthcare industry collects enormous amounts of healthcare data and required to mine and ascertain hidden information for constructive decision making. In recent years, the computer technology and machine learning methods with data mining approach increase techniques in assisting the doctors for productive decision making related to heart disease and stroke identification at an early stage. The needs to reduce the pattern matching loss by performing pattern matching while effectively improving the heart disease identification at much early stage poses severe challenges to the database community. In this paper a method called Clinical Heart Disease-Decision Supportive Optimized Mining (CHD-DSOM) is presented to overcome the pattern matching loss, and perform perfect pattern matching. The method CHD-DSOM categories to enable decision support with multi-dimensional analysis using Mahalanobis distance measure for obtaining dynamic data table information and typically built to support early stage of stroke identification and levels of heart disease. This helps in identifying the solution for different patterns and therefore reducing the
pattern matching loss. With the objective of identifying the level of heart disease in a very accurate manner, the CHD-DSOM method uses the Iterative Dichotomiser 3 based decision tree model. With Iterative Dichotomiser 3 based decision tree model, the stroke is also identified very easily by starting with the original set S as the root node. The entropy value of every attribute is calculated and is placed in the decision tree. Decision tree are constructed from the higher to lower values to perform easy pattern matching, aiming at reducing the processing time for pattern matching. Experiment is conducted with the Cleveland Clinic Foundation Heart disease data set available from UCI repository using the factors such as pattern matching loss rate, accuracy, processing time for pattern matching, computational cost. Experimental analysis shows that the CHD-DSOM method is able to reduce the processing time for pattern matching by 33.23% and reduce the pattern matching loss rate by 29.67% compared to the state-of-the-art works.

References

3. Jae-Gil Lee, Jiawei Han, Xiaolei Li and Hong Cheng, “Mining Discriminative Patterns for Classifying Trajectories on Road Networks”, IEEE Transactions on Knowledge and Data Engineering, Volume 23, Issue 5, May 2011, Pages 713 – 726.
8. Weiguo Zheng, Lei Zou, Xiang Lian, Dong Wang and Dongyan Zhao, “Efficient Graph Similarity Search Over Large Graph Databases”, IEEE Transactions on Knowledge and Data Engineering, Volume 27, Issue 4, April 2015, Pages 964 – 978.
Index Terms

Computer Science  Information Sciences

Keywords

Machine Learning, Decision Supportive, Optimized Mining, Pattern Matching, Decision Tree Iterative Dichotomiser