Abstract

Diabetes mellitus or simply diabetes is a disease caused due to the increase level of blood glucose. Various available traditional methods for diagnosing diabetes are based on physical and chemical tests. These methods can have errors due to different uncertainties. A number of Data mining algorithms were designed to overcome these uncertainties. Among these algorithms, amalgam KNN and ANFIS provides higher classification accuracy than the existing approaches. The main data mining algorithms discussed in this paper are EM algorithm, KNN algorithm, K-means algorithm, amalgam KNN algorithm and ANFIS algorithm. EM algorithm is the expectation-maximization algorithm used for sampling, to determine and maximize the expectation in successive iteration cycles. KNN algorithm is used for classifying the objects and used to predict the labels based on some closest training examples in the feature space. K-means algorithm follows partitioning methods based on some input parameters on the datasets of n objects. Amalgam combines both the features of KNN and K means with some additional processing. ANFIS is the Adaptive Neuro Fuzzy Inference System which combines the features of adaptive neural network and Fuzzy Inference System. The data set chosen for classification and experimental simulation is based on Pima Indian Diabetic Set from University of California, Irvine (UCI) Repository of Machine Learning databases.
References

- UCI machine learning repository and archive. ics. uci. edu/ml/datasets.html
Index Terms

Computer Science
Data Mining

Keywords

Data mining Diabetes EM algorithm KNN algorithm K-means algorithm amalgam KNN algorithm ANFIS algorithm