Abstract

A new hybrid data mining model is proposed to provide a comprehensive analytic method for finding an optimal number of different pathological types of any disease and its complications, an optimal partitioning representative and extracts the most significant features for each pathological type. This model is an integration of both characteristics of supervised and...
An Effective Intelligent Model for Finding an Optimal Number of Different Pathological Types of Diseases

unsupervised models and is based on clustering, feature selection, and classification concepts. This model takes into consideration access to the highest classification accuracy during the clustering process. Experiments have been conducted on 3 real medical datasets related to the diagnosis of breast cancer, heart disease, and post-operative infections. The performance of this method is evaluated using information entropy, squared error, classification sensitivity, specificity, overall accuracy, and Matthew's correlation coefficient. The results show that the highest classification performance is obtained using our proposed model, and this is very promising compared to NaïveBayes, Linear Support Vector Machine (Linear SVM), Polykernal Support Vector Machine (Polykernal SVM), Artificial Neural Network (ANN), and Support Feature Machines (SFM) models.

References

An Effective Intelligent Model for Finding an Optimal Number of Different Pathological Types of Diseases


Index Terms

Computer Science

Data Mining
An Effective Intelligent Model for Finding an Optimal Number of Different Pathological Types of Diseases

**Keywords**

Clustered feature selection classification

SFM model

- breast cancer
- heart disease
- post-operative infection