Abstract

In this work, ANN and SVM, two most popular supervised machine learning techniques, are considered as the representatives and k-means clustering is used as representative of unsupervised learning. By analyzing the diagnosis result using Wisconsin Breast Cancer Dataset (WBCD) which is commonly used among researchers who use machine learning methods for breast cancer diagnosis, it can be concluded that SVM outperforms in case of breast cancer diagnosis. The result is verified using two other breast cancer datasets. One is Breast Cancer Dataset from UCI Machine Learning Repository and another one is "Breast cancer dataset with Electrical Impedance Measurements in samples of freshly excised tissue". The purpose of the comparison is to choose the best solution in terms of performance. Another notable significance of the work is that accuracy of the recognition drops down severely if proper feature set is not used. One significant disadvantage of neural network
is its time taken to build the model which is also evident from the work.

References


Dataset: http://archive.ics.uci.edu/ml/datasets/breast+tissue

Index Terms

Computer Science
Applied Sciences

Keywords
Ann  SVM  K-means Clustering