

عنوان مقاله:

Feature Subset Selection and Parameters Optimization for Support Vector Machine in Breast Cancer Diagnosis

محل انتشار:

دوازدهمین کنفرانس ملی سیستم های هوشمند ایران (سال: 1392)

تعداد صفحات اصل مقاله: 6

نویسندگان:

Elnaz Olfati - Department of Electrical Engineering Imam Khomeini International University Qazvin, Iran

Hassan Zarabadipour - Department of Electrical Engineering Imam Khomeini International University Qazvin, Iran

Mahdi Aliyari Shoorehdeli - Department of Mechatronics Engineering K. N. Toosi University of Technology Tehran, Iran

خلاصه مقاله:

Due to high death rate in women with breast cancer, the detection will play a major role in the treatment of this type of cancer. Therefore, the early detection of breast cancer willincrease the patients' chances of survival. The main tendency in feature extraction has been illustrating the data in a lower dimensional and different feature space, for instance, using principal component analysis (PCA). In this paper, we argue that feature selection depend on top of eigenvalue certainly is notproper because they may not encode useful information for classi1cation purposes, features should be selected form all the components by feature selection methods. So, Genetic Algorithm (GA) is used in the most favorable selection of principal components instead of using classical method. We have applied PCA for dimension reduction, genetic algorithms for featureselection and support vector machines for classification. Theestimate of this Algorithm has been done based on Wisconsin Breast Cancer Dataset (WBCD) which is commonly used amongresearchers who use machine learning methods for breast cancer diagnosis. The performance of this approach is given. In addition, the methods used in the past have been compared to the performance of the chosen approach. This approach affordsoptimal classification which is capable to minimize amount of features and maximize the accuracy sensitivity, specificity and receiver operating characteristic (ROC) curves. 10-fold crossvalidationhas been used on the classification phase. The average classification accuracy of the developed PCA+GA+SVM system isobtained 100% for a subset that contained two features. This is very favorable compared to the previously reported .results

کلمات کلیدی:

component; Breast cancer diagnosis; Principal component analysis (PCA); Genetic algorithm (GA); Support vector machine(SVM); Feature subset selection

لینک ثابت مقاله در پایگاه سیویلیکا:

https://civilica.com/doc/276346

