

Correlation-based feature selection for association rule mining in semantic annotation of mammographic medical images

ABSTRACT

Mining of high dimension data for mammogram image classification is highly challenging. Feature reduction using subset selection plays enormous significance in the field of image mining to reduce the complexity of image mining process. This paper aims at investigating an improved image mining technique to enhance the automatic and semi-automatic semantic image annotation of mammography images using multivariate filters, which is the Correlation-based Feature Selection (CFS). This feature selection method is then applied onto two association rules mining methods, the Apriori and a modified genetic association rule mining technique, the GARM, to classify mammography images into their pathological labels. The findings show that the classification accuracy is improved with the use of CFS in both Apriori and GARM mining techniques.

Keyword: Correlation-based feature selection; Multivariate filters; Association rule mining; Mammographic image classification; Semantic annotation