

Using fuzzy association rule mining in cancer classification

ABSTRACT

The classification of the cancer tumors based on gene expression profiles has been extensively studied in numbers of studies. A wide variety of cancer datasets have been implemented by the various methods of gene selection and classification to identify the behavior of the genes in tumors and find the relationships between them and outcome of diseases. Interpretability of the model, which is developed by fuzzy rules and linguistic variables in this study, has been rarely considered. In addition, creating a fuzzy classifier with high performance in classification that uses a subset of significant genes which have been selected by different types of gene selection methods is another goal of this study. A new algorithm has been developed to identify the fuzzy rules and significant genes based on fuzzy association rule mining. At first, different subset of genes which have been selected by different methods, were used to generate primary fuzzy classifiers separately and then proposed algorithm was implemented to mix the genes which have been associated in the primary classifiers and generate a new classifier. The results show that fuzzy classifier can classify the tumors with high performance while presenting the relationships between the genes by linguistic variables.

Keyword: Breast cancer; Fuzzy classifier; Gene expression profile; Gene selection