Abstract

In fuzzy theory, there's something called clustering or grouping. Clustering is data processing to classes or clusters so that the data in a class has high similarity with the others which happen to be in the same class, but it is different with the data in the other class. Fuzzy clustering is a technique to determine the optimal cluster in a vector that is based on the normal Euclidian for the distance between vectors. One of the cases that related to the data grouping is a patient's tumor grouping into two classes: benign and malignant. The growth of the tumor can lead to a benign tumor or the malignant one. The tumor in breast tumor that is a malignant growth can be called breast cancer.
Fuzzy C-Means algorithm is one of the fuzzy clustering algorithms that was used the most because it tries to cluster every elements that is the member of a cluster to a cluster. Fuzzy C-Means algorithm can be used to cluster data that was known the number of the cluster. First of all, the initial process is to do preprocessing data (outside the system) using weka tools, that is the attribute selection process. This process is done to get the selected attribute that will be used as inputs for the system. Then data will be processed with fuzzy c-means algorithm. The last prosess to be done is calculating the system's accuracy. The testing result shows that the fuzzy c-means method is capable to identificate the breast cancer risk with the good accuracy.

Keywords: breast cancer, fuzzy clustering, fuzzy c-means.