ABSTRACT

Tuberculosis (TBC) is an infectious disease that causes the second largest health problem in the world after HIV, so it is still a global concern. The mortality and morbidity rate caused by mycobacterium tuberculosis is also high. In this study, the author will apply the data mining classification method, namely the Naïve Bayes method and the K-Means clustering method. The Naïve Bayes method is used to assist in the early diagnosis of tuberculosis in order to reduce the spread of the disease to the community. Meanwhile, K-Means clustering was used to determine the pattern of the spread of tuberculosis that occurred in Padangsidimpuan City. This research is how to develop Geographic Information System (GIS) as a distribution control tool for government agencies in Padangsidimpuan City. Based on the results of the K-Means clustering, it is found that cluster 1 (green) consists of 9 members of health care facilities, cluster 2 (yellow) consists of 4 members of health care facilities, and cluster 3 (red) consists of 2 members of health care facilities. The application of the nave Bayes method for the TB disease diagnosis system by means of training symptom data and then calculating the selected symptoms to find the probability value of positive or negative TB patients, the highest result is taken as a diagnosis result.

Keywords: Geographic Information System (GIS), K-Means, Naïve Bayes, TBC