

**Abstract**

Cancer or commonly known by the people of Indonesia as a tumor disease is a disease that results from abnormal growth in body tissue cells. Cancer is included in the category of dangerous diseases that can cause death. Based on World Health Organization data - WHO, in 2015 no less than 8.8 million human birth due to cancer. The number is increasing every year. Even so, 30 to 50% of these cases can be prevented.

Therefore the program is to remove the death rate. Collecting patient data containing information about patients with cancer is the first step in mortality rates. Microarray DNA (gene expression) has many things in the context of cancer.

The research was conducted using Minimum Redundancy Maximum Relevance (mRMR) algorithm to select number of features, and Probabilistic Neural Network (PNN) algorithm to classify. With the feature selection method used, obtained 10 best features. Of these 10 features can be classified types of illnesses suffered by patients based on the probability of distance between features. Accuracy obtained from PNN classification algorithm of 90.91%. Gene expression data are leukemia data (blood cancer) obtained from Kent Ridge Biomedical Data set Repository.

**Keywords: data mining, classification, minimum redundancy maximum relevance, probabilistic neural network**