
Abstract

News is one means of information for the general public. Today, the number of news articles that reach 2 million articles per day can make it difficult for users to find news articles they want to read. In order to make it easier for users, most Indonesian newspapers classify their articles into certain categories, but there are also many blogs, or amateur articles that have not classified the news they circulated. Therefore this paper aims to categorize Indonesian language news using the weighted k-nearest neighbor method. In this paper there are several stages in classifying the news, namely preprocessing, feature extraction, and classification using wK-NN. The study used the wK-NN method where $K = 6$. In this study feature extraction was carried out in *unigram* and *bigram* which resulted in accuracy that was not much different. So it is recommended to use *unigram* because it is more efficient

Keywords: Classification, Preprocessing, feature extraction, weighted K-nearest neighbor
