

Abstract

Needs for news information has increased since the change from physical media to online media. News is grouped according to categories to making it easier for readers to get the news as desired. Grouping to determine the category of news information is known as text classification. The number of words in the news text create diversity of words that appear and can be minimized by the stemming process, which is changing an affixed word into its root word. This study comparing between use of stemming and without stemming, find the best value of K and optimum distance calculation of K-Nearest Neighbor. The best results were obtained with the classification conditions without applying stemming algorithm with number of K=9 with cosine distance for distance calculation which resulted accuracy in 0,9671. This result is greater than the classification that applies stemming algorithm in condition K=7 using cosine distance which resulted accuracy in 0,9660.

Keywords: news, text classification, k-nearest neighbor, *stemming*