Abstract

The development of information technology at this time is increasing rapidly. The amount of data continues to grow as more and more people are utilizing technological conveniences as a place to find information, one of which is news. The number of news raises issues in classifying every news topic. Therefore, a system is required that can classify every news topic in the class it should be. This research aims to create a system that can categorize the news on the class it should be. The method of Weighted K-Nearest Neighbor is used as a classification method because it is a simple classifier. News datasets are retrieved from the Kaggle website with five news classifications which is Business, Entertainment, Politics, Sports, and Tech. In this research there are several stages in the system, namely preprocessing data, extraction of features and classification using Weighted K-Nearest Neighbor. Based on the test result the best accuracy obtained is equal to 95,9% where the k-neighbor value = 34.

Keywords: classification, preprocessing, feature extraction, weighted k-NN