

Abstract

Children's books are one source of knowledge for readers, especially children. When the book is read, the child will try to make sense of every word and sentence in it. There was a problem when a content error was found in the book. The content in question is words and sentences that have meanings that are not polite, sexual, and rude words. For children at the elementary school level, the content becomes meaningful reflectivity (taboo). Based on these problems, a final assignment research was carried out on children's stories taken from fiction books and textbooks. This research was conducted by building a system for detecting reflector content on story text that is used as a dataset. The study was conducted by building a model using the k-Nearest Neighbor text classification algorithm with a distance measure approach. Distance measure used is Euclidean Distance and Manhattan Distance. The system is evaluated using precision, recall, and F1 score. Based on the evaluation results, testing scenarios using Euclidean distance and Manhattan distance get a precision value of 0.915, recall 0.845, and F1 score 0.895.

Keywords: children's book, distance measure, gereflekter, k-Nearest Neighbor