

Abstract

Textbooks and storybooks are the ones used as a source of knowledge. When children read a book, they will try to interpret each word and sentence in it. However, it will be a problem if the book contains vulgar words and indecent sentences. For children at the elementary school level, it is not allowed. For this research, we called that content as gereflekted content. Based on these problems, this research was conducted by building a system to detect gereflekted content in the text of the child's stories that were used as a data set. A system is built by using Naïve Bayes Classifier (NBC) and then evaluated in two scenarios using accuracy, precision, and recall metrics because the characteristics of the data set are imbalanced with the amount of data in the negative class are greater than the data in the positive class. From evaluation results, test scenario produced a high average precision of 99.01%, whereas the recall value has an average of above 50%. From these two values, it can be concluded that the model built by the system has not detected the class properly, but highly trusted when it does.

Keywords: children's story, gereflekted content, Naïve Bayes Classifier