Abstract

Product review text is a text that contains a review, rating or review of a work. There are many types of product review texts, one of which is a beauty product review text that serves as input to improve the quality of a beauty product. However, many beauty product review texts use informal language, so text normalization is necessary. Text normalization is a term used to convert text formats with a predetermined purpose. The use of text normalization at the pre-processing stage is usually able to improve the quality of the text of the beauty product reviews to be classified. In this final project, text normalization is implemented for beauty product reviews, with the aim of improving the performance of sentiment classification. Text normalization is carried out using a dictionary-based method, which is defined based on a list of internet slang and spelling correction. The experimental results show that with dictionary-based text normalization, there is an accuracy difference of 0.54% between the classification process without text normalization and the classification process with text normalization.

Keywords : Text normalization, Beauty product review text, Internet slang, Spelling correction