

Aspect and Opinion Extraction of Indonesian Lipsticks Product Reviews using Conditional Random Field (CRF)

Deffri Kun Indarta¹, Ade Romadhony²

^{1,2} School of Computing, Telkom University, Bandung

¹deffrikunindarta@students.telkomuniversity.ac.id, ²aderomadhony@telkomuniversity.ac.id

Abstract

Aspect and opinion extraction is a key process in several downstream applications, such as market analysis. The aspect and opinion extraction of lipsticks product reviews in Indonesia is modeled using Conditional Random Field (CRF) method. Product reviews contain important information, consists of aspects and opinions which are very influential for customers to make their decisions towards the products. However, with many product review opinions available on social media and e-commerce, many opinions and aspects that can be obtained in a product review, analyzing, and extracting information aspect and opinion becomes progressively difficult and time-consuming. The dataset contains review text written not only in standard and colloquial Indonesian languages but also standard and colloquial English, labeled by BIO format notation. The experimental results show the average F1 score on 10 aspect opinion labels is 44.1% with accuracy of 81.8%. The results on the baseline method, HMM, show lower F1 and accuracy scores. The errors found in the results have mostly occurred when the same words have different meanings. This error has a percentage of 75% of the total result. Second largest error caused by unknown words with the percentage of 14% of the total result.

Keywords : product review, aspect extraction, opinion extraction, Conditional Random Field.
