Abstract

Nowadays, there are merchants selling their product on the web. They usually ask their customer to give a review about product they have purchased. It will make another potential customer being easier to make a decision for buying a product they will bought. The amounts of the review will make the potential customer difficult for reading each of the review. Therefore, this thesis aims to mining an opinion on the review.

Process mining in this thesis is called opinion mining or sentiment analysis. Where the output was to determine whether a sentence is a sentence positive or negative opinions. This thesis oriented feature-based opinion in which subjective orientation is determined, prior to first extract the product entities commented. The first process is to extract features and their opinions are followed, then extract subjective sentences, and finally determine the orientation of opinion (positive or negative). The method used in this system is Lexicalized Hidden Markov Model (L-HMM) that integrates linguistic features, such as part-of-speech, into an automated machine learning. The system is expected to be able to recognize the new vocabulary trainer has not shown up in the data, which will be implemented together with the smoothing technique. In obtaining the data train and its features, opinion and orientation in large quantities, do bootstrapping techniques.

The test results show that given bootstrapping techniques can be used to enrich the data used directly trainer. In addition, the use of L-HMM capable of extracting features and their orientation with good opinions in a sentence.

Keywords **:** *data mining, sentiment analysis, feature-based opinion, part-of-speech, lexicalized hidden markov models* (L-HMM)