

Abstract

At this time, many online forums have sprung up, especially those that discuss product reviews. From the written product reviews, we can extract some important information, one of which is the mention of the product name. Extracting product names in a text has several benefits, such as market intelligence management, enterprise content management, and others. However, because review texts are usually in the form of informal texts, variations in the mention of a product name vary, so to extract product names with a definite match to the list of brand names is very difficult. Therefore, a product named entity recognition (PRO NER) recognition system can be established to overcome this problem. In building the PRO NER system in this final project, the Hidden Markov Model (HMM) method and Viterbi algorithm are used. Based on the HMM-Viterbi model that has been built, the obtained F1-scores for each brand (BRA), product (PRO), and type (TYP) labels are 69.17%, 44.65%, dan 26.67%.

Keywords: NER, information extraction, PRO NER, Hidden Markov Model, Viterbi algorithm