

We employed a semi-supervised learning approach on corpus development. The corpus development process consists of two phases: initial manual labeling and automatic labeling. The initial manual labeling was performed on 2,275 sentences, while the automatic labeling was performed on 16,853 sentences. We employed Conditional Random Field (CRF) classifier on automatic labeling process, which used labeled training dataset from the initial manual labeling process output. The experimental result shows the corpus accuracy is 73.66%.