

Abstract

Javanese is the one of regional languages that is used mostly on the Indonesian island of Java. Used in the provinces of Central Java, DIY, East Java, and surrounding areas. Using data from online news in the Javanese language. The method used is CRF (Conditional Random Fields). CRF is a structure for building probabilistic models. This paper describes the part of speech for Javanese using Conditional Random Fields (CRF) method. We used a corpus of size 3000 words and cross-validation for training. Our model gave an overall accuracy of 67%.

Keywords—Part of Speech, Conditional Random Fields