

ABSTRAK

POS *tagger* adalah sebuah *software* yang berfungsi untuk memberikan label gramatika pada setiap kata dalam sebuah teks. Pembuatan POS *tagger* Bahasa Jawa sudah dilakukan oleh beberapa peneliti. Namun, akses *resource* dari POS *tagger* yang mereka buat tidak terbuka. Tujuan dari penelitian ini adalah membuat POS *tagger* yang memiliki *resource* dengan akses terbuka dan memiliki antarmuka yang ramah pengguna. Pada riset ini, pembuatan POS *tagger* Bahasa Jawa dilakukan menggunakan aplikasi Nooj, dengan pendekatan *rule-based* atau linguistik. Pendekatan *rule-based* dalam pembuatan *tagger* adalah pendekatan yang membutuhkan rumus, leksikon, dan konfigurasi. Tahapan pembuatan *tagger* yakni pembuatan kamus, konfigurasi, *testbed*, dan pendataan distribusi afiksasi. Hasil penelitian menunjukkan 1) pembuatan POS *tagger* verba Bahasa Jawa membutuhkan tiga instrumen; kamus, tata bahasa (rumus), dan konfigurasi, 2) tingkat akurasi *tagger* yang Peneliti buat adalah 86.3%, dan 3) distribusi produktivitas afiksasi yang paling produktif dalam Bahasa Jawa pada korpus Javanese UD adalah afiksasi dengan jenis prefiks. Pada riset ini, peneliti telah menunjukkan bahwa pembuatan POS *tagger* Bahasa Jawa, selain menggunakan pendekatan probabilitas seperti di ilmu komputer, juga dapat dilakukan menggunakan pendekatan linguistik. *Tagger* verba ini dapat dikembangkan agar mencakup semua kelas kata dalam Bahasa Jawa.

Kata kunci: *tagger*, verba, afiksasi, dan Bahasa Jawa

ABSTRACT

POS tagger is a software that functions to provide grammatical labels for each word in a text. The creation of a Javanese POS tagger has been conducted by several researchers. However, the resource access of the POS tagger they created is not open. The objective of this research is to create a Javanese POS tagger that features open-access resources and a user-friendly interface. The tagger was created using the NooJ application and employs a rule-based or linguistic approach. This approach relies on rules, a lexicon, and specific configurations. The creation process involved dictionary creation, configuration, testbed setup, and the documentation of affixation distribution. The results of the study show 1) the creation of a Javanese verb POS tagger requires three instruments: dictionary, grammar (formula), and configuration, 2) the accuracy level of the tagger I created is 86.3%, and 3) the distribution of the most productive affixation productivity in the Javanese UD corpus is affixation with the prefix type. In this study, the researcher have shown that creating a Javanese POS tagger, in addition to using a probabilistic approach like in computer science, can also be done using a linguistic approach. This verb tagger can be extended to cover all word classes in Javanese.

Keywords: tagger, verb, affixation, and Javanese