

Implementasi *Dependency Parser* Menggunakan Metode Jaringan Saraf Tiruan

Nurul Izzah¹, Moch Arif Bijaksana², Arief Fatchul Huda³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹nurulizzah@students.telkomuniversity.ac.id, ²arifbijaksana@telkomuniversity.ac.id,

³afhuda@uinsgd.ac.id

Abstrak

Beberapa tahun terakhir, *parsing* menjadi sangat populer dalam ruang lingkup NLP (*Natural Language Processing*) dengan adanya *Dependency Parser*. Namun, hampir semua *Dependency Parser* yang ada melakukan klasifikasi berdasarkan jutaan fitur indikator jarang (*sparse*). Fitur ini tidak hanya buruk dalam mengambil kesimpulan, tetapi juga membatasi kecepatan parsing secara signifikan sehingga parsing yang dihasilkan tidak maksimal. Untuk mengatasi permasalahan tersebut, dilakukan pergantian penggunaan fitur jarang (*sparse*) menjadi fitur padat (*dense*). Dengan penggunaan fitur padat dapat secara efektif mengurangi *sparsity* antar kata. Metode klasifikasi Jaringan Saraf Tiruan digunakan untuk menghasilkan *parsing* yang cepat dan ringkas dalam *Transition-Based Dependency Parser* dengan menggunakan 2 *hyperparameter*. Dataset yang digunakan pada penelitian ini yaitu Bahasa Arab, China, Inggris, dan Indonesia. Berdasarkan evaluasi yang telah dilakukan, menunjukkan hasil yang lebih tinggi dengan menggunakan *hyperparameter* kedua. Pada pengujian dengan data *test* Bahasa Inggris diperoleh nilai akurasi LAS (*Labeled Attachment Score*) 80.4% dan UAS (*Unlabelled Attachment Score*) 83%. Kemudian dengan data *dev* diperoleh nilai akurasi LAS 81.1% dan UAS 83.7%, serta kecepatan *parsing* sebesar 98 *sent/s*.

Kata kunci : Parsing, dependency parser, transition-based dependency parser.

Abstract

In recent years, parsing has become very popular within the scope of NLP (Natural Language Processing) with the Dependency Parser. However, almost all existing Dependency Parser do classifications based on millions of sparse indicator features. This feature is not only bad in drawing conclusions, but also significantly limits the speed of parsing so that the resulting parsing is not optimal. To overcome these problems, changing the use of sparse features becomes dense features to reduce sparsity between words. The Artificial Neural Network classification method is used to produce fast and concise parsing in the Transition-Based Dependency Parser by using 2 hyperparameters. The dataset used in this study is Arabic, Chinese, English, and Indonesian. Based on the evaluation that has been done, it shows a higher result using the second hyperparameter. In testing with English test data, the accuracy value of LAS (Labeled Attachment Score) is 80.4% and UAS (Unlabelled Attachment Score) is 83%, Then with dev data obtained an accuracy value of LAS 81.1% and UAS 83.7%, and parsing speed of 98 sent/s.

Keywords: Parsing, dependency parser, transition-based dependency parsing.
