
Abstrak

Berita merupakan alat yang digunakan untuk menyebarkan informasi melalui berbagai media, salah satunya adalah internet. Berbagai macam artikel berita memiliki kata-kata yang tidak dikenal dalam kamus seperti kata-kata gaul dan kata-kata asing yang tidak ada dalam korpus. Bagaimana model POS *tagging* yang dibangun di atas korpus dapat menangani pelabelan kelas kata dalam berita bahasa Indonesia. Penelitian ini dilakukan untuk mengetahui hasil POS *tagging* pada kumpulan berita tentang Telkom University yang diseleksi secara manual. Dengan menggunakan model *bidirectional LSTM*, tiga skenario pengujian dicoba untuk meningkatkan kinerja model yang dibangun, skenario pengujian menerapkan *padding* terbaik untuk korpus, membandingkan hasil kinerja model korpus yang dimodifikasi dengan model korpus asli, dan menentukan dimensi vektor Word2vec. Kemudian model terpilih dari masing-masing korpus diimplementasikan pada berita yang telah diberi label secara manual. Salah satu pengujian skenario terbaik diperoleh dengan memodifikasi korpus dengan menghilangkan kata rangkap pada kelas kata "X" dan mengubah beberapa kelas kata "X" yang lebih cenderung merupakan kata asing sehingga diubah menjadi kelas kata "FW". Hasil kinerja terbaik dalam implementasi pemberitaan *Telkom University* menggunakan model *bidirectional LSTM* yang dibangun berdasarkan korpus modifikasi mendapatkan nilai *accuracy* sebesar 92.74%, *precision* sebesar 92.85%, *recall* sebesar 92.74%, dan *f1-score* sebesar 92.48%.

Kata kunci : POS *tagging*, *bidirectional LSTM*, Bahasa Indonesia, berita
