

ABSTRAK

PERFORMA MODEL NAMED ENTITY RECOGNITION MENGUNAKAN METODE BLSTM-HMM PADA DOKUMEN TWITTER BERBAHASA INDONESIA

Oleh

Adi Wijaya

19102099

Twitter merupakan platform *microblogging* yang populer di Indonesia, menyediakan data teks dalam bahasa Indonesia yang dapat dimanfaatkan untuk aplikasi *Natural Language Processing (NLP)*. *Named Entity Recognition (NER)* adalah sub-tugas penting dalam *Information Extraction (IE)* yang bertujuan untuk mengidentifikasi entitas bernama seperti orang, lokasi, dan organisasi dalam teks. Keterbatasan dalam penelitian *NER* bahasa Indonesia di platform Twitter menimbulkan urgensi untuk mengembangkan model *NER* yang dioptimalkan. *NER* memiliki peran penting dalam *IE* dan *NLP* secara umum, memungkinkan ekstraksi informasi yang berguna secara otomatis dari teks Twitter. Adapun tujuan dari penelitian ini yaitu, untuk mengetahui nilai performa *F1 score* pada model *NER* menggunakan metode *BLSTM-HMM*, dengan fitur *FastText* dan *POS tag*. Serta mengetahui apakah penerapan metode tersebut bisa meningkatkan performa *F1 score* terhadap model *NER*. Penelitian ini mengimplementasikan metode *Bidirectional Long Short-Term Memory (BLSTM) – Hidden Markov Model (HMM)* dengan fitur *FastText* dan *POS tag* untuk meningkatkan kinerja model *NER* pada data dokumen twitter berbahasa indonesia. Model yang dikembangkan berhasil mencapai *F1 score* sebesar 89,23%. Penggabungan model *NER* dengan *BLSTM-HMM* serta penambahan fitur *FastText*, dan *POS tag* dapat menjadi alternatif yang menarik dalam meningkatkan kinerja pengenalan entitas. Dalam penelitian ini, penggabungan tersebut berhasil memberikan hasil yang positif.

Kata kunci: *BLSTM, FastText, HMM, Named Entity Recognition, POS tag*