

ABSTRAK

Terciptanya internet, jejaring sosial, forum, dan teknologi informasi yang tersebar secara cepat, menyebabkan interaksi terhadap informasi semakin sulit untuk dipahami, dibuat, dikembangkan, dan disimpan. Dengan luasnya informasi sehingga hampir tidak mungkin untuk seorang pun untuk memproses dan meringkas semua data informasi yang tersedia. Indonesia memiliki literasi yang sangat rendah dari negara lain dengan beberapa faktor seperti belum membiasakan membaca buku dari rumah, perkembangan teknologi yang semakin pesat, sarana membaca yang minim, kurangnya motivasi untuk membaca, dan malas untuk mengembangkan gagasan. *Automatic text summarization* adalah salah satu alternatif teknologi yang bisa digunakan untuk menyelesaikan masalah tersebut. *Automatic text summarization* merupakan cabang dari ilmu *Natural Language Processing* (NLP) yang bertujuan untuk merepresentasikan dokumen teks panjang yang dikompresi sehingga informasi yang didapatkan dengan cepat dimengerti dan dapat dibaca oleh pengguna. Telah dilakukan berbagai metode untuk mengatasi masalah *automatic text summarization* untuk objek Bahasa Indonesia yaitu berbasis *extractive* dan *abstractive*. Untuk mengatasi masalah ini, pada penelitian digunakan *extractive text summarization* berbasis *machine learning*. Pada penelitian ini menggunakan *dataset* publik yang bisa digunakan untuk penelitian-penelitian selanjutnya. Metode yang digunakan untuk mendapatkan hasil *summarization* dengan menggunakan metode *Word2Vec* dengan penerapan model *Continuous Bag-of-Word* (CBOW) dan *Skip-Gram*. Metode yang digunakan untuk evaluasi akurasi hasil ringkasan adalah *Recall-Oriented Understudy for Gisting Evaluation* (ROUGE).

Kata Kunci: *Automatic Text Summarization, Word2Vec, Continuous Bag-of-Words, Skip-Gram, Recall-Oriented Understudy for Gisting Evaluation*