

INTISARI

Peringkasan Ekstraktif Otomatis pada Berita Berbahasa Indonesia Menggunakan A Hierarchical Structured Self-Attentive Model for Extractive Document Summarization (HSSAS)

Oleh

Nathanael Gavin Hasiholan Siregar

17/412646/PA/17965

Dengan semakin banyaknya masyarakat yang berpindah dari media cetak ke media digital dan berita daring yang tersedia di internet, maka dibutuhkan suatu sistem peringkasan otomatis untuk artikel berbahasa Indonesia. Sistem ini akan mempermudah masyarakat yang ingin membaca banyak berita untuk dapat mengetahui dan memahami intisari dari keseluruhan berita dengan cara yang mudah dan dalam waktu yang singkat tanpa harus membaca seluruh isinya.

Penelitian ini mengimplementasikan metode peringkasan artikel otomatis A Hierarchical Structured Self-Attentive Model for Extractive Document Summarization (HSSAS) pada *dataset* artikel berbahasa Indonesia dan mengevaluasi performa dari metode tersebut dengan menggunakan metode ROUGE untuk membandingkan kualitas ringkasan yang dibuat secara otomatis dengan ringkasan yang dibuat manual oleh manusia yang tersedia pada *dataset* INDOSUM.

Pengujian dilakukan dengan menghitung skor *f-measure* dari metrik evaluasi ROUGE-1, ROUGE-2, dan ROUGE-L. Berdasarkan pengujian yang telah dilakukan, metode HSSAS memperoleh skor *f-measure* ROUGE-1 sebesar 67.37%, ROUGE-2 sebesar 60.65%, dan ROUGE-L sebesar 66.61%.

Kata Kunci: Peringkasan Teks Otomatis, Ekstraktif, HSSAS

ABSTRACT

Automatic Extractive Summarization for Indonesian News Articles Using A Hierarchical Structured Self-Attentive Model for Extractive Document Summarization (HSSAS)

By

Nathanael Gavin Hasiholan Siregar

17/412646/PA/17965

As many people are beginning to migrate from printed newspaper to the online version, the necessity of an automatic summarization for Indonesian news articles is also increasing. This system will help everyone who wants to read a lot of information from news articles to know and understand the essence of every news article easily and in a relatively short time without having to read the whole news article.

This research implements A Hierarchical Structured Self-Attentive Model for Extractive Document Summarization (HSSAS) model to process a dataset that consists of Indonesian news articles. Then, this research evaluates the performance of the model using the ROUGE method to compare the quality of the summary that is generated automatically by the model with the one that is arranged manually by humans, which is available on the dataset.

The evaluation is done by calculating the f-measure score of the metric evaluation ROUGE-1, ROUGE-2 and ROUGE-L. Based on the result of the evaluation, the HSSAS model get f-measure score 67.37% for ROUGE-1, 60.65% for ROUGE-2, and 66.61% for ROUGE-L.

Keywords: Automatic Text Summarization, Extractive, HSSAS