

INTISARI

PENGARUH VARIATIONAL AUTOENCODER TERHADAP RINGKASAN EKSTRAKTIF TEXT RANK

Oleh

Laksita Kusuma Wardhani

21/476211/PPA/06155

Dataset dalam dunia digital memiliki berbagai macam karakteristik, baik dari ukuran, tema, jenis, dan lain sebagainya, dataset dalam dunia digital memiliki variasi yang luas, hal tersebut juga diiringi oleh jumlah data melimpah, data yang melimpah ini memiliki masalah nya sendiri, hal ini dikarenakan ukuran dimensional vektor nya akan menjadi sangat besar untuk diolah, bahkan untuk memuat beberapa data, bisa menghabiskan waktu sehari penuh, hal tersebut juga memengaruhi kinerja dari peringkasan teks yang akan dilakukan selanjutnya

Berdasarkan masalah diatas, maka diusulkan sebuah solusi dimana dimensional data dapat dikurangi ukurannya, yaitu dengan menerapkan *Variational Autoencoder*, model ini yang akan membantu merubah data menjadi lebih ringan dengan *value* yang berbeda dari awal sebelumnya, selanjutnya untuk peringkasan teks otomatis, digunakan peringkasan ekstraktif, spesifiknya *TextRank*, peringkasan ini dipilih dikarenakan waktu komputasi nya yang lebih cepat dibandingkan ringkasan abstraktif.

Untuk mengetahui nilai kualitas ringkasan yang baik, maka akan diimplementasikan alat ROUGE, alat ini digunakan dengan cara mengurangi hypothesis, yang merupakan peringkasan buatan sistem, serta reference, ringkasan dari dataset. Dengan adanya alat ini dapat dihasilkan bahwa ROUGE *TextRank* baik yang melalui *Variational Autoencoder* ataupun tidak memiliki nilai yang tidak baik atau dibawah 0.5, untuk hasil ROUGE paling baik diperoleh *text preprocessing* dengan *Non-Stemming* serta waktu komputasi tercepat didapatkan oleh *text preprocessing Stemming*.

Kata Kunci : *Variational Autoencoder*, *TextRank*, ROUGE

ABSTRACT

IMPACT OF VARIATIONAL AUTOENCODER TOWARDS EXTRACTIVE SUMMARIZATION TEXT RANK

By

Laksita Kusuma Wardhani

21/476211/PPA/06155

Datasets in the digital world have various characteristics, both in terms of size, theme, type, and so on, datasets in the digital world have wide variations, this is also accompanied by an abundant amount of data, this abundant data has its own problems, this is because the size of the dimensional vector will be very large to process, even to load some data, it can take a whole day, this also affects the performance of the text summarization that will be carried out next

Based on the above problems, a solution is proposed where the dimensional data can be reduced in size, namely by implementing Variational Autoencoder, this model will help change the data to be lighter with different values from the previous beginning, then for automatic text summarization, extractive summarization is used, specifically TextRank, this summarization is chosen because its computing time is faster than abstractive summaries.

To find out the value of a good summary quality, the ROUGE tool will be implemented, this tool is used by reducing the hypothesis, which is a system-made summary, and reference, a summary of the dataset. With this tool, it can be shown that ROUGE TextRank, whether through Variational Autoencoder or not, has a poor value or below 0.5. The best ROUGE results are obtained by text preprocessing with Non-Stemming and the fastest computing time is obtained by text preprocessing Stemming.

Keywords : *Variational Autoencoder, TextRank, ROUGE*



UNIVERSITAS
GADJAH MADA

PENGARUH VARIATIONAL AUTOENCODER TERHADAP RINGKASAN EKSTRAKTIF TEXTRANK

Laksita Kusuma Wardhani, Yunita Sari, S.Kom., M.Sc., Ph.D.

Universitas Gadjah Mada, 2024 | Diunduh dari <http://etd.repository.ugm.ac.id/>