**Peringkasan Artikel Berbahasa Indonesia Menggunakan *TextRank* dengan Pembobotan BM25**

**Yurdha Fadhila Hernawan 1, Putra Pandu Adikara 2, Randy Cahya Wihandika 3**

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya

Email: 1penulis.satu@xmail.ac.id, 2penulis.dua@xmail.ac.id

**Abstrak**

Tempatkan abstrak berbahasa Indonesia pada bagian ini. Abstrak memberikan gambaran umum tentang isi makalah dan harus ditulis dengan *Times New Roman* 11 dalam format satu kolom. Panjang ideal sebuah abstrak adalah 150 sampai 200 kata.

**Kata kunci**: *kata kunci sedapat mungkin menjelaskan isi tulisan, ditulis dengan huruf kecil kecuali singkatan, maksimum enam kata, masing-masing dipisahkan dengan koma, Times New Roman 10, italic*

**Abstract**

*Tempatkan abstrak berbahasa Inggris pada bagian ini. Gunakan font Times New Roman 11, italic.*

**Keywords**: *kata kunci sedapat mungkin menjelaskan isi tulisan, ditulis dengan huruf kecil kecuali singkatan, maksimum enam kata, masing-masing dipisahkan dengan koma, Times New Roman 10, italic*

# PENDAHULUAN

Penggunaan internet sebagai sumber informasi telah membawa manusia pada era *one click away*. Apa pun bisa diakses di mana pun kapan pun, baik secara visual maupun tidak. Namun, apakah setiap informasi yang diakses selalu sesuai dengan konteks yang diinginkan? Bisa dikatakan hanya sedikit pengguna yang dapat memahami semua informasi ketika membaca sebuah tulisan panjang (Niu et al., 2016). Kesulitan tersebut akan membuat pengguna untuk membaca ulang, sehingga akan menghabiskan banyak waktu.

Untuk memudahkan pengguna internet dalam mendapatkan informasi yang ringkas dengan tidak merusak atau menghilangkan informasi penting, maka dibutuhkan suatu peringkasan otomatis (Abbasi-ghalehtaki et al., 2016). Berdasarkan metode yang digunakan, peringkasan teks dapat dikategorikan dalam dua bentuk, yaitu ekstraktif dan abstraktif. Ringkasan ekstraktif merupakan ringkasan yang terdiri atas kumpulan dari bagian-bagian penting suatu tulisan yang dapat mewakili keseluruhan teks, sedangkan ringkasan abstraktif merupakan ringkasan yang terdiri dari kalimat baru yang dapat merepresentasikan konteks tulisan dalam bentuk lain. Selain itu, peringkasan teks juga dapat dikelompokkan berdasarkan dokumen yang digunakan menjadi *single document* dan *multi-document* (Fang et al., 2017).

Masalah utama yang muncul setelah melakukan peringkasan adalah kualitas hasil peringkasan. Apakah konteks yang dibicarakan pada hasil ringkasan sudah dapat merepresentasikan tulisan secara utuh. Penelitian oleh Pinandhita (2013) telah melakukan peringkasan ekstraktif teks berbahasa Indonesia dengan menggunakan kata benda yang terdapat dalam sebuah dokumen. Penelitian tersebut menghitung nilai kemiripan (*similarity*) antara kalimat berdasarkan kata benda yang terdapat pada setiap kalimat, lalu memeringkatkan kalimat yang paling penting dengan mengurutkan total nilai kemiripan tersebut. Evaluasi dilakukan dengan membandingkan hasil kinerja beberapa metode *similarity* agar dapat menentukan metode *similarity* terbaik dalam melakukan peringkasan.

Selain itu, peneliti lain telah melakukan penelitian untuk mendapatkan ringkasan secara abstraktif (Niu et al., 2016). Penelitian tersebut menggunakan dokumen teks opini pendek berbahasa China dengan mengelompokkan teks yang mirip lalu meringkasnya. Proses pengelompokan teks yang mirip dilakukan dengan menggunakan metode *K-Means* dengan fitur yang didapatkan dari nilai *word2vec* (nilai hubungan kedekatan antara kata). Hasil pengelompokan teks akan diperingkatkan menggunakan metode *TextRank*. Kalimat yang dianggap penting dan dapat dijadikan ringkasan memiliki peringkat tertinggi. Hasil pemeringkatan akan dijadikan ringkasan dengan menggunakan *encoder-decoder Reccurent Neural Network* (RNN) untuk membentuk kalimat baru.

Penelitian sebelumnya yang menjadi acuan dalam penelitian ini adalah peringkasan esktraktif menggunakan data *statistical* (Mussina, Aubakirov and Trigo, 2018). Data statistical merupakan nilai kesamaan *text units* dalam sebuah dokumen. *Text units* dalam dokumen dapat berupa kata, kalimat atau paragraf. Penelitian tersebut menggunakan metode *TextRank* untuk merepresentasikan kalimat dalam sebuah dokumen dan menggunakan beberapa metode *similarity*.

Berdasarkan pemaparan di atas, penulis mengajukan penelitian untuk melakukan peringkasan otomatis dengan objek artikel berita online berbahasa Indonesia menggunakan metode pemeringkatan *TextRank* dan BM25 sebagai fungsi similarity. BM25 dipilih berdasarkan penelitian sebelumnya yang telah melakukan pengujian menggunakan beberapa fungsi *similarity* dan mendapatkan bahwa BM25 menghasilkan ringkasan yang lebih baik dari pada fungsi *similarity* lainnya (Barrios et al., 2016).

# METODOLOGI PENELITIAN

Persamaan matematika harus diberi nomor urut dalam kurung biasa dan harus diacu dalam tulisan.

Persamaan matematika dinomori dengan angka Arab di dalam tanda kurung buka-tutup pada posisi rata kanan kolom. Persamaan ditulis menjorok ke dalam sejauh 6 mm.

Untuk persamaan yang tidak cukup ditulis dalam lebar 1 kolom, penulisannya dapat melintasi 2 kolom, ditulis di bagian bawah halaman dan diberi nomor urut yang sesuai. Perhatikan contoh penulisan Persamaan (1).

(1)

# TABEL DAN GAMBAR

Tabel dan gambar harus diberi nomor dan judul lengkap serta harus diacu dalam tulisan. Contoh: Tabel 1, Tabel 2(a), Gambar 1, Gambar 2(a). Berikut ini diberikan contoh penulisan tabel dan gambar.

## 3.1. Penulisan Tabel

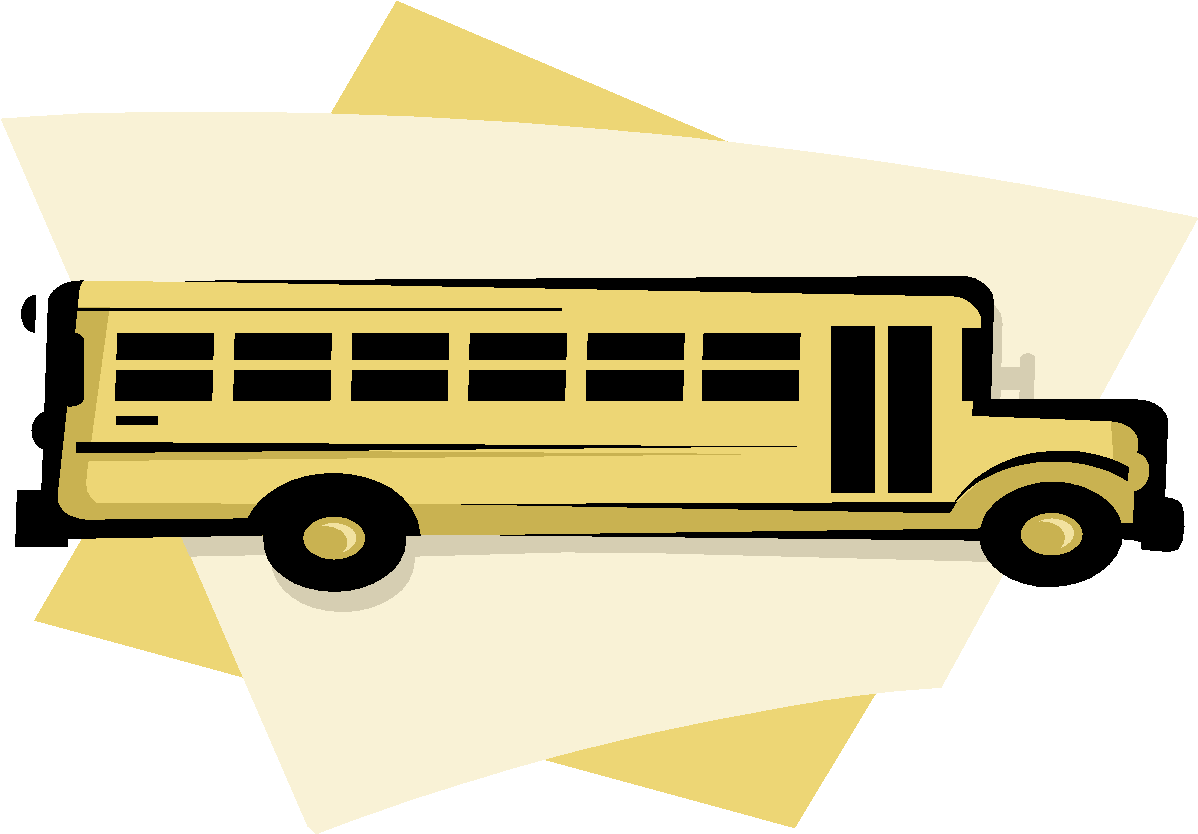
Perhatikan contoh penulisan Tabel 1. Keterangan tabel dituliskan pada bagian atas tabel. Sebisa mungkin hindari penggunaan garis vertikal.

Tabel 1. Rancangan Analisis Komputasi

|  |  |  |
| --- | --- | --- |
| **mesin** | **operasi** | **waktu(menit)** |
| 1 | 5 | 10 |
|  | 4 | 12 |
|  | 4 | 16 |
| 2 | 3 | 18 |
|  | 4 | 20 |

## 3.2. Penggunaan Gambar

Perhatikan contoh penggunaan Gambar 1. Keterangan gambar dituliskan pada bagian bawah gambar.



Gambar 1. Contoh penggunaan gambar

# SUMBER PUSTAKA/RUJUKAN

Sumber pustaka/rujukan sedapat mungkin merupakan pustaka-pustaka terbitan 10 tahun terakhir. Pustaka yang diutamakan adalah sumber-sumber primer berupa laporan penelitian (termasuk Skripsi/Tugas Akhir, Tesis, Disertasi) atau naskah-naskah penelitian dalam jurnal dan/atau majalah ilmiah.

Cara penulisan di daftar pustaka bisa dilihat pada contoh Daftar Pustaka. Sumber pustaka dituliskan terurut alfabetis dan kronologis.

Sumber pustaka yang digunakan bisa berupa:

* makalah jurnal;
* makalah konferensi ilmiah (*proceeding*);
* buku teks;
* laporan penelitian;
* skripsi atau thesis;
* makalah dalam buku kumpulan makalah ilmiah (*book section*).

# ATURAN LAIN

Semua naskah ditelaah secara blind-review oleh mitra bestari (reviewers) yang ditunjuk oleh redaksi menurut bidang kepakarannya. Penulis naskah diberi kesempatan untuk melakukan perbaikan (revisi) naskah atas dasar rekomendasi/saran dari mitra bestari dan redaksi pelaksana. Kepastian pemuatan atau penolakan naskah akan diberitahukan secara tertulis melalui email.

Pemeriksaan dan penyuntingan cetak-coba dikerjakan oleh redaksi dan/atau dengan melibatkan penulis. Naskah yang sudah dalam bentuk cetak-coba dapat dibatalkan pemuatannya oleh redaksi jika diketahui bermasalah.

Segala sesuatu yang menyangkut perijinan pengutipan atau penggunaan software komputer untuk pembuatan naskah atau hal lain yang terkait dengan HAKI yang dilakukan oleh penulis naskah, berikut konsekuensi hukum yang mungkin timbul karenanya, menjadi tanggung jawab penuh penulis naskah tersebut.

# CONTOH DAFTAR PUSTAKA

Dao, S. D. & Marian, R. 2011. Optimisation of precedence-constrained production sequencing and scheduling using genetic algorithms. *Proceedings of the International Multi Conference of Engineers and Computer Scientists*, 16-18 March, Hong Kong.

Gen, M. & Cheng, R. 2000. *Genetic Algorithms and Engineering Optimization.* John Wiley & Sons, Inc., New York.

Liliana, D. Y. & Mahmudy, W. F. 2006. Penerapan Algoritma Genetika pada Otomatisasi Penjadwalan Kuliah. *Laporan Penelitian DPP/SPP.* FMIPA Universitas Brawijaya, Malang.

Marian, R. M., Luong, L. & Dao, S. D. 2012. Hybrid genetic algorithm optimisation of distribution networks—a comparative study. *Dalam:* AO, S. I., CASTILLO, O. & HUANG, X. (editor.) *Intelligent Control and Innovative Computing.* Springer, US.

Phanden, R. K., Jain, A. & Verma, R. 2013. An approach for integration of process planning and scheduling. *International Journal of Computer Integrated Manufacturing*, 26(4), 284-302.

Ridok, A. 2014. Peringkasan dokumen Bahasa Indonesia berbasis non-negative matrix factorization. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, 1(1), 39-44.

Tala, F. Z. 2003. A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia. *Ph.D. Thesis*. Universiteit van Amsterdam.

Wang, L. 2007. *Process planning and scheduling for distributed manufacturing.* Springer, London.

Wibawa, A. P., Nafalski, A. & Mahmudy, W. F. 2013. Javanese `speech levels machine translation: improved parallel text alignment based on impossible pair limitation. *IEEE International Conference on Computational Intelligence and Cybernetics*, 3-4 December, Yogyakarta, Indonesia. 16-20.