

Deteksi Konten Berita Hoax Tentang Dunia Kesehatan Dimedia Sosial Dengan Menggunakan Metode Text Rank

Mulyadi Syahputra Napitupulu

Program Studi Teknik Informatika, Fakultas Ilmu Komputer Dan Teknologi Informasi, Universitas Budi Darma, Medan, Indonesia
Jl. Sisingamangaraja No. 338, Medan, Sumatera Utara, Indonesia
Email: mulyadisyahtputra855@gmail.com

Abstrak—Dengan menggunakan Metode Text Rank, diharap kita dapat mengklasifikasikan berita mengenai hal yang sedang beredar dikalangan masyarakat khususnya berita yang sedang viral akhir-akhir ini mengenai virus corona di media sosial seperti Facebook dan media sosial berbentuk text lainnya. Disini penulis memberikan gambaran bagi pembaca yakni masyarakat agar cerdas dan tanggap terhadap berita yang disebarakan tersebut, sehingga dapat mengurangi kecemasan masyarakat mengenai berita tersebut. penulis juga ingin memperjelas berita tersebut kepada pembaca yakni masyarakat agar lebih bijak menerima serta memilah berita-berita yang disebarakan sehingga masyarakat tidak terpropokasi dengan berita-berita dari orang-orang yang tidak bertanggung jawab dan hanya ingin memperkeruh suasana saja. Telah terjadi berbagaimacam kasus penyebaran berita yang bukan merupakan fakta atau sering disebut hoax. Sedangkan hoax adalah informasi berbahaya yang menyesatkan persepsi manusia dengan menyebarkan informasi yang salah namun dianggap sebagai kebenaran. Media komunikasi sangat penting untuk menyebarkan informasi atau berita yang sedang terjadi namun banyak media komunikasi yang digunakan untuk menyebarkan berita simpang siur yang belum tentu dapat terbukti kebenarannya tersebar di seluruh media komunikasi yang dapat membuat orang yang membaca mempercayai berita bohong tersebut.

Kata Kunci: metode, text rank, hoaks, berita, berbentuk text

Abstract— By using the Text Rank Method, it is hoped that we can classify news regarding things that are currently circulating among the public, especially news that has gone viral recently regarding the corona virus on social media such as Facebook and other text-based social media. Here the author provides an illustration for readers, namely the public, to be intelligent and responsive to the news that is distributed, so that it can reduce public anxiety about the news. The author also wants to clarify the news to readers, namely the public, so that they are wiser in accepting and sorting out the news that is distributed so that the public is not promoted by news from irresponsible people who only want to muddy the waters. There have been various cases of spreading news that is not fact or is often called a hoax. Meanwhile, hoaxes are dangerous information that misleads human perception by spreading false information but is considered to be the truth. Communication media is very important for disseminating information or news that is currently happening, but many communication media are used to spread confusing news which cannot necessarily be proven to be true, spread throughout communication media which can make people who read believe the fake news.

Keywords: Method, text, rank, hoax, news, in text form

1. PENDAHULUAN

Dalam situasi dunia saat ini yang dilanda oleh wabah Covid-19 atau yang paling dikenal sebagai Virus Corona sangatlah merisaukan masyarakat, bukan hanya dinegara kita saja Indonesia bahkan diseluruh warga Negara saat ini, pada penyajian informasi yang uptodate dan tepat guna dapat membantu dimasa saat ini, media sosial merupakan salah satu wadah yang sangat mendukung untuk penyebaran guna membantu dan mempererat komunikasi antara sesama manusia [1][2][3].

Pada saat ini tiada yang bisa menjadi penghalang untuk saling berbagi komunikasi. Tidak hanya sebagai media hiburan, media sosial juga bisa berperan sebagai media penyebaran informasi. Informasi yang tersebar seseorang maupun kelompok melalui media sosial akan dengan cepat dicerna oleh setiap lini masyarakat luas. Berbagi informasi antar masyarakat merupakan hal yang positif, namun dizaman saat ini tidak seluruh informasi yang disebarakan melalui media sosial berupa informasi yang fakta.

Berbagaimacam kasus telah terjadi dan penyebaran beritanya pun sudah tersebar baik dari mulut kemulut maupun melalui media sosial. Hal tersebut memungkinkan bahwa berita yang real tersebut ditambah oleh beberapa kalangan sehingga berita yang fakta berubah menjadi bukan fakta atau sering disebut hoax. Media komunikasi sangat penting untuk menyebarkan informasi atau berita yang sedang terjadi namun banyak media komunikasi yang digunakan untuk menyebarkan berita yang berisi argument dari beberapa pihak saja. Tidak sedikit masalah yang ditimbulkan oleh tersebarnya berita hoax di kalangan masyarakat, sebagai contoh di bulan Maret, 2020 tersebar berita bawang dapat menjadi anti virus corona (Covid-19). Setelah beberapa pihak meneliti berita dan melakukan ujicoba tersebut, virus covid-19 tidak bisa disembuhkan bahkan mencegah virus tersebut [4][5][6].

Melalui hal ini mahasiswa dihimbau untuk memiliki kepedulian sosial terhadap berita atau informasi yang disebarakan. Mahasiswa diajak untuk menjadi generasi yang kritis terhadap suatu informasi, bukan hanya diterima mentah mentah. Mahasiswa juga diharap dapat memilah mana berita yang hanya *real* dan dapat di sebarakan secara langsung dan mana berita yang perlu diubah sesuai dengan berita yang sesuai dengan berita sesungguhnya agar tidak membuat kecemasan yang meluas di masyarakat [7][8].

Banyak hal terbukti bahwa tersebarnya berita yang berisi argumen dari beberapa pihak atau hoax dapat menimbulkan permasalahan yang besar, banyak orang yang melebih-lebihkan fakta tentang suatu hal. Masyarakat Indonesia termasuk masyarakat yang kurang peduli tentang apa yang mereka bagikan bahkan masyarakat Indonesia

dapat membagikan berita yang salah karena tidak membaca dengan tuntas suatu informasi yang diterima melalui media sosial maupun aplikasi pesan singkat yang berbasis text, termasuk juga kalangan terdidik seperti mahasiswa [9][10].

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahapan yang akan dilakukan dalam pengerjaan penelitian ini adalah sebagai berikut:

- a. Studi Pustaka
Pemahaman studi pustaka tentang konsep dan Metode *Text Rank* dalam Deteksi Konten Berita Hoax Tentang Dunia Kesehatan Dimedia Sosial, Dengan Metode tersebut kita bisa menggunakan bahasa pemrograman *visual C#* yang berintegrasi dengan *database access*
- b. Analisis dan Perancangan Aplikasi
Pada tahap ini dilakukan analisis dengan metode *Text Rank* yang digunakan dalam mendeteksi konten berita Hoax
- c. Dokumentasi (*Documentation*)
Menyusun berkas penelitian berdasarkan penelitian dan hasil yang telah dilakukan sesuai dengan sistematika penelitian

2.2 Berita

Berita adalah kata yang sering kita dengar dalam kehidupan kita sehari-hari, berita juga sebagai salah satu cara untuk mendapatkan sebuah informasi mengenai suatu kejadian yang sudah atau sedang terjadi, dikarenakan berita sangat mudah untuk diakses karena berita disajikan lewat bentuk cetakan, siaran, internet, atau dari mulut ke mulut kepada orang ketiga sehingga mudah untuk didapatkan. Menurut Uskali berita sebagai media memiliki peran fungsi sebagai jembatan komunikasi antar berbagai perbedaan budaya dan pola pikir yang menyebar di masyarakat [11][12][13]. Saat ini pemberitaan di seluruh dunia semakin rumit atas berbagai kompleks masalah yang semakin melebar mulai dari isu sara, hoax dan lain sebagainya, maka media berita dituntut untuk semakin bekerja secara profesional dengan menjaga kualitas berita yang disampaikan. Berikut adalah beberapa syarat agar berita yang disampaikan terjaga kualitasnya:

- a. Berita harus merupakan fakta dan aktual (kejadian terkini).
- b. Berita harus seimbang (tidak memihak).
- c. Berita harus lengkap (mengandung semua unsur berita).
- d. Berita harus menarik dan bermanfaat.
- e. Berita harus disusun dengan sistematis.

Sedangkan dalam penyusunan sebuah berita maka perlu diperhatikan unsur yang harus dipenuhi agar berita tersebut dapat dikatakan lengkap. Syarat berita dikatakan lengkap adalah jika sebuah memiliki 5 W + 1 H, yaitu:

- a. *What*, yaitu berita harus memuat mengenai apa yang sedang terjadi.
- b. *Who*, yaitu berita harus memuat informasi mengenai siapa saja yang terlibat dalam kejadian tersebut.
- c. *Why*, yaitu berita harus memuat latar belakang kejadian tersebut terjadi.
- d. *When*, yaitu berita harus memuat informasi mengenai kapan yang diberikan terjadi.
- e. *Where*, yaitu berita harus memberikan informasi tempat dan lokasi kejadian.
- f. *How*, yaitu berita harus memuat mengenai bagaimana kejadian tersebut bisa terjadi.

2.3 TextRank

TextRank merupakan algoritma berbasis graf bekerja dengan cara memberikan peringkat pada teks dengan cara merepresentasikan objek dalam teks (Mihalcea & Tarau, 2004). Dengan algoritma ini mengandalkan bentuk graf maka untuk dapat menentukan sebuah kalimat penting atau tidak digunakan simpul (*node/verteks*). Dalam menerapkan algoritma ini, langkah pertama yang dilakukan adalah membangun sebuah graf yang terdiri dari titik (*verteks*) mewakili tiap kalimat. Setiap kalimat akan dihubungkan berdasarkan kesamaan (*similarity*) antar kalimat biasa disebut dengan *edge*. *TextRank* juga merupakan suatu peringkasan atau pemeringkatan teks yang sangat memudahkan bagi setiap orang yang ingin menggunakannya baik dalam dokumen dan sebagainya. Dimana kebutuhan akan informasi di saat ini sangat penting dan kebutuhan akan informasi tersebut dapat diperoleh dari berbagai media, seperti halnya media digital atau pun media cetak. Untuk dapat mengetahui informasi penting dari suatu dokumen atau teks, pembaca harus banyak meluangkan waktu apalagi ketika sedang membaca berita yang lagi booming saat ini seperti halnya berita tentang virus corona, pembaca pasti akan banyak menemukan kata yang sering muncul pada setiap paragraph berita [14].

Oleh sebab itu, jika salah satu berita tersebut di media sosial terlebih dalam bentuk text akan diproses secara cepat dan dapat diringkas oleh suatu sistem tanpa menghilangkan informasi yang penting, maka pembaca dapat menghemat waktu. Karena dengan begitu pembaca dapat memahami dan mengetahui informasi tersebut adalah berita yang *real* adanya ataukah hanya sebuah prediksi dari seseorang yang akan meresahkan masyarakat luas. Meringkas teks pada berita terutama dalam membahas tentang virus secara manual oleh manusia, memerlukan waktu yang lama, apalagi teks berita banyak dan memakai bahasa kesehatan tersebut banyak dan panjang oleh. Karena itu, diperlukanlah sistem peringkasan otomatis untuk mengatasi banyaknya waktu. Pada penelitian ini akan dibuat sistem peringkasan teks

otomatis yang menggunakan algoritma *TextRank* sebagai metodenya dan juga dibuat berbasis database atau rancang aplikasi sehingga dapat digunakan oleh siapa saja yang memerlukannya.

TextRank adalah suatu metode yang digunakan untuk membantu pengguna dalam mencari suatu informasi yang berguna dalam kumpulan besar teks. Dalam *TextRank*, pencarian representasi suatu teks pada kamus dikenal dengan *indexing*. Proses dari Algoritma *TextRank* atau *indexing* ini meliputi beberapa tahapan yaitu tahap text cleaning, case folding, tokenization, stopword removal (eliminasi stopword), stemming, dan term *weight* [15]. Sebagai contoh dari metode yang digunakan yaitu dengan melakukan beberapa tahapan atau tahap representasi dibawah ini dengan menggunakan rumus perhitungan yaitu :

Tabel 1. Tahap Representasi Text Rank

Vertex 1	Vertex Perbandingan	
kalangan menilai resesi ekonomi	orang superkaya dunia memprediksi resesi 2020	Vertex 1
	menghadapi resesi.	Vertex 2
	orang superkaya menghadapi penurunan ekonomi memindahkan aset berisiko	Vertex 3
	risiko rendah	
	aset dimanfaatkan orang orang kaya resesi	Vertex 4
	peneliti institute for development of economics and finance bhima yudhistira aset	Vertex 5
	dimanfaatkan orang superkaya deposito surat utang negara jepang bertenor jangka pendek	
	persiapkan dana simpanan aset relatif aman deposito rupiah 2 miliar dijamin	Vertex 6
	lembaga penjamin simpanan surat utang pemerintah jepang dijamin negara tenor pendek bhima dihubungi detikcom	

$$\text{Similarity}(S_i, S_j) = \frac{|W_k| \ W_k \in S_i \ \& \ W_k \in S_j}{\log(|S_i|) + \log(|S_j|)} \dots \dots \dots 1$$

$$\text{Similarity}(S_1, S_3) = \frac{1}{4 + 10} \dots \dots \dots 2$$

$$(1/14 = 0.07)$$

Keterangan rumus W_k = Jumlah kata (term) yang sama antara kalimat S_i dan S_j , S_i = Panjang kalimat S_i , S_j = Panjang kalimat S_j . Menurut Anwar (2015) keunggulan dari algoritma *TextRank* antara lain:

- Unsupervised*, artinya algoritma ini bekerja tanpa perlu data training untuk memproses data.
- Language independent*, Dikarenakan *TextRank* hanya mengandalkan kata, kalimat ataupun paragraf yang terdapat dalam teks sehingga pengetahuan mendalam akan sebuah bahasa tidak begitu diandalkan.

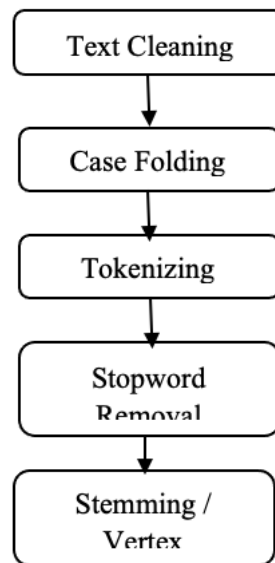
3. HASIL DAN PEMBAHASAN

Dengan menggunakan metode *TextRank*, maka dengan mudah kita melakukan eliminasi guna mengetahui konten berita yang sedang beredar dengan penggunaan *Text Rank*, hal ini dapat dilihat dimana dan bagaimana kata-kata yang sering muncul ditemukan. Dengan demikian, sebagai contoh adalah “Berita pagi ini yang menghebohkan dunia”, dari kata menghebohkan. Orang mengartikan kata tersebut sebagai kata yang serius dan menanggapi sebagai hal yang amat serius saat itu. Hal itu belum tentu benar, diharapkan pembaca mengetahui bahwa banyak kata menghebohkan yang pembaca temukan didalam setiap konten berita ketika sedang membacanya, namun ketika pembaca ingin mengetahui dimana saja letak kata tersebut berada, pembaca tidak bisa begitu cepat mengetahui dimana kata tersebut dan berapa jumlah kata yang sama dan sering muncul dalam suatu konten berita..

3.1 Penerapan Metode Text Rank

Menerapkan metode *TextRank* untuk menyusun suatu Konten berita diindonesia maka dilakukan dengan membandingkan penggunaan kata-kata yang berbeda dari kata dasarnya yang sama, mencari frekuensi dari kata dalam dokumentasi, mencari kata-kata kunci, membuat indeks dan daftar kata yang digunakan dalam Berita dan membantu untuk mengetahui kebenaran konten berita sehingga tidak ada yang salah paham dengan arah konten berita yang tersebar saat itu. Berdasarkan penelitian yang didapatkan sering kali ketika membaca konten berita, pembaca banyak menemukan kata yang sering muncul pada setiap kalimat ataupun setiap paragrafnya yang terkadang tidak dipahami oleh pembaca arah kata tersebut, sehingga menimbulkan rasa penasaran kepada pembaca.

TextRank merupakan algoritma peringkat berbasis grafik untuk memproses teks yang telah dikelompokkan berdasarkan nilai yang sering muncul dan dengan *TextRank* juga membantu pengelompokkan kata kata yang berulang dan mengidentifikasi berita yang dianggap hoax. Adapun langkah atau tahap proses yang dilakukann dalam penyelesaian untuk melakukan penomoran atau mempermudah pencarian kata yang sering muncul yaitu:



Gambar 1. Algoritma TextRank

Dengan tahap – tahap inilah akan dilakukannya proses untuk mempermudah pencarian pada hasil kata yang sering muncul dan yang sudah dihitung. Adapun beberapa tahap – tahapan yang akan dilakukan dengan menggunakan metode TextRank ini antara lain sebagai berikut:

a. Tahap pertama (*Text Cleaning*)

Melakukan *Text Cleaning* yang bertujuan untuk menghilangkan setiap tanda baca yang tidak penting pada kata-kata disetiap kalimat pada konten berita yang akan diringkas untuk mengetahui setiap kata yang sering muncul, tanda baca yang dimaksud ialah seperti tanda petik dua, petik satu dan lainnya kecuali tanda titik.

b. Tahap kedua (*Case Folding*)

Setelah tanda baca yang tidak penting dihilangkan maka dilakukan *Case Folding* yang bertujuan untuk menormalkan huruf atau mengubah setiap besar atau kapital menjadi huruf kecil.

c. Tahap ketiga (*Tokenizing*)

Selanjutnya dilakukan tahap *Tokenizing* yang bertujuan untuk membentuk satu-satu kalimat menjadi kalimat yang lebih jelas dan tepat.

d. Tahap keempat (*Stopword removal*)

Setelah dilakukannya *Tokenizing* maka selanjutnya dilakukan tahap *Stopword Removal* dimana akan dilakukannya pemilihan kata pada konten berita yang tidak penting atau tidak perlu sehingga lebih mudah untuk menemukan kata yang sering muncul.

e. Tahap kelima (*Stemming* atau *Vertex Term*)

Setelah semua tahap dilakukan maka dilanjutkan dengan membuat sebuah vertex yang berupa kolom dan baris yang berisi setiap kalimat isi konten berita, vertex ini bertujuan untuk menghitung jumlah kata yang sama atau sering muncul pada setiap kolom dan baris pada vertex. Jumlah vertex ditentukan seberapa banyak kalimat atau paragraph yang ada pada text konten berita. Setelah semua kata yang terdapat pada vertex diketahui jumlahnya maka dilakukanlah penjumlahan dengan menggunakan rumus yang ada atau yang sudah ditentukan, dengan demikian jika hasil sudah diketahui maka dapat dilakukan sebuah peringkatan, dimana hasil yang lebih tinggi dan rendah dapat dikelompokkan dengan menggunakan atau membuat sebuah grafik / graf untuk mengetahui setiap jumlah tertinggi dan terendah. Setelah sudah dilakukannya peringkatan dan sudah mengetahui setiap jumlah kata tertinggi dan terendah maka selanjutnya dilakukanlah penomoran dengan menggunakan Strong Concordance, penomoran dilakukan pada setiap kata yang terbanyak berupa kata yang sama atau pun kata yang sering muncul.

4. KESIMPULAN

Dari pembahasan sebelumnya penulis dapat membuat beberapa kesimpulan yang dimana dalam penggunaan metode *Textrank* untuk deteksi konten berita Hoax. Adapun kesimpulan yang dapat diperoleh yaitu Dalam membaca berita yang tersebar dimedia sosial, setidaknya kita seharusnya memahami kriteria dari berita agar berita yang kita baca tersebut tidak mempengaruhi tidakan kita dimasyarakat atau dunia nyata dan tidak mudah terpengaruh oleh berita hoax. Mampu mengimplementasikan perhitungan metode *textrank*, dalam perhitungan text dari berita yang diuji, sehingga hasil berita yang sedang diuji mendapatkan kesimpulan utama dari berita-berita tersebut. Dengan aplikasi yang dirancang diharapkan kegiatan dalam perhitungan agar mempermudah medeteksi konten berita sehingga dapat hasil yakni kesimpulan utama berita. atau dapat menyebarkan informasi bagi kepentingan public.

REFERENCES

- [1] A. W. Fathurrahman, M. Thoriqulhaq, And F. Arianto, “Penerapan Machine Learning Untuk Pengklasifikasian Hoaks Pada Platform Media Sosial,” *Senada*, Vol. 2022, No. Senada, Pp. 66–68, 2022, [Online]. Available: <https://Senada.Upnjatim.Ac.Id/Index.Php/Senada/Article/View/48%0ahttps://Senada.Upnjatim.Ac.Id/Index.Php/Senada/Article/Download/48/26>.
- [2] R. Wati, “Penerapan Algoritma Naive Bayes Dan Particle Swarm Optimization Untuk Klasifikasi Berita Hoax Pada Media Sosial,” *Jitk (Jurnal Ilmu Pengetah. Dan Teknol. Komputer)*, Vol. 5, No. 2, Pp. 159–164, 2020, Doi: 10.33480/Jitk.V5i2.1034.
- [3] E. Edward, “Identifikasi Berita Hoax Berbasis Web Menggunakan Algoritma C4.5,” *J. Ilmu Komput. Dan Sist. Inf.*, Vol. 9, No. 1, P. 53, 2021, Doi: 10.24912/Jiksi.V9i1.11558.
- [4] Y. T. Harahap, “Perbandingan Kompresi Data Teks Dengan Metode Lzy, Lzap Dan Lzw.”
- [5] M. F. Arfabuma, “Implementasi Algoritma Textrank Dalam Automatic Text Summarization Pada Buku Digital Berbahasa Indonesia Untuk Meningkatkan Kognitif Pembaca/M. Fandi Arfabuma.” Universitas Negeri Malang, 2021.
- [6] Y. T. Handika, S. Defit, And G. W. Nurcahyo, “Text Mining Dalam Membandingkan Metode Naive Bayes Dengan C.45 Dalam Mengidentifikasi Berita Hoax Pada Media Sosial,” *Rang Tek. J.*, Vol. 5, No. 1, Pp. 116–123, 2022, Doi: 10.31869/Rtj.V5i1.2855.
- [7] M. A. Zamzam, “Sistem Automatic Text Summarization Menggunakan Algoritma Textrank.” Universitas Islam Negeri Maulana Malik Ibrahim, 2019.
- [8] A. Y. S. A. Setiawan And E. Alexander, “Pengujian Algoritma Textrank Dalam Merangkum Teks,” In *Prosiding Seminar Nasional Amikom Surakarta*, 2023, Pp. 25–35.
- [9] I. D. R. Riadi, L. Wikarsa, And S. Pandelaki, “Implementasi Algoritma Textrank Untuk Peringkasan Dokumen.” Universitas Katolik De La Salle Manado, 2022.
- [10] E. Eris And J. Pragantha, “Penerapan Algoritma Textrank Untuk Automatic Summarization Pada Dokumen Berbahasa Indonesia,” *J. Ilmu Tek. Dan Komput.*, Vol. 1, No. 1, Pp. 71–78, 2017.
- [11] B. Kurniawan, S. Effendi, And O. S. Sitompul, “Klasifikasi Konten Berita Dengan Metode Text Mining,” *J. Dunia Teknol. Inf.*, Vol. 1, No. 1, Pp. 14–19, 2012.
- [12] I. R. Munthe, B. H. Rambe, R. Pane, D. Irmayani, And M. Nasution, “Sistem Pendukung Keputusan Pemilihan Pimpinan Dengan Metode Multi Attribute Utility Theory (Maut) Di Pt. Sagami Indonesia,” *J. Mantik*, Vol. 3, No. January, Pp. 31–38, 2019.
- [13] D. Marta, G. L. Ginting, And A. M. H. Sihite, “Deteksi Berita Palsu Tentang Vaksinasi Covid-19 Dengan Menggunakan Text Mining Dan Algoritma Cosine Similarity,” *Komik (Konferensi ...)*, Vol. 6, No. November, Pp. 129–139, 2023, Doi: 10.30865/Komik.V6i1.5738.
- [14] C. Khontoro, J. Andjarwirawan, And Y. Yulia, “Penerapan Algoritma Textrank Dan Dice Similarity Untuk Verifikasi Berita Hoax,” *J. Infra*, Vol. 9, No. 1, Pp. 98–102, 2021.
- [15] A. Y. S. A. Setiawan And E. Alexander, “Penerapan Algoritma Textrank Dalam Merangkum Teks Word Dan Pdf.”