

Analisa Sentimen Informasi Hoaks Pasca Pandemi Covid-19 dengan Text Mining

Akhmad Rezki Purnajaya¹, Yonky Fernando^{2,*}

¹Fakultas Komputer, Program Studi Teknik Perangkat Lunak, Universitas Universal, Batam, Indonesia

²Fakultas Komputer, Program Studi Teknik Informatika, Universitas Universal, Batam, Indonesia

Email: ¹rezkipurnajaya@email.com, ^{2,*}yongkyfernando194@gmail.com

Email Penulis Korespondensi: yongkyfernando194@gmail.com

Abstrak—Pada era globalisasi ini, penyampaian informasi dapat disampaikan dengan media informasi apapun, hal ini juga yang dapat menyebabkan terjadinya penyebaran informasi hoaks. Hal ini sangat merugikan bagi penerima informasi, terkhususnya pada informasi yang sangat penting untuk diterima, salah satu contohnya adalah informasi terkait virus Covid-19 atau Corona yang telah melanda dunia maupun Indonesia. Berdasarkan data dari Gugus Tugas Percepatan Penanganan Covid-19, di era pasca pandemi masih ditemukan 227 informasi hoaks yang telah diidentifikasi. Maka dari itu, pada penelitian ini dilakukan penelitian informasi hoaks, terkhususnya tentang virus Covid-19 pada era pasca pandemi menggunakan teknik Text Mining untuk mengetahui pola informasi yang sering muncul pada pasca pandemi Covid-19. Dalam makalah ini juga akan menunjukkan beberapa hasil tentang keterkaitan beberapa kata dengan kata utama pada informasi hoaks di era pasca pandemi Covid-19 dalam bentuk Term of Frequency, Word Cloud, Layout Fruchterman Reingold, dan Layout Circle. Hasil dari penelitian ini yaitu pada era pasca pandemi Covid-19, informasi hoaks seringkali berhubungan dengan ‘vaksin’ dan ‘covid’, dengan masing-masing sebanyak 113 dan 111 kata. Hal itu dikarenakan masyarakat yang meragukan keamanan dan efektivitas vaksin Covid-19 yang menyebabkan penyebaran informasi hoaks untuk menghindari vaksin tersebut. Kata-kata lain yang sering muncul dalam informasi hoaks pasca pandemi adalah ‘omicron’, ‘divaksin’, ‘varian’, ‘pfizer’, ‘mrna’, ‘virus’, dan ‘vaksinasi’, dengan jumlah kata yang lebih rendah.

Kata Kunci: Covid-19; Text Mining; Hoaks; Pasca Pandemi; Analisa Sentimen

Abstract—In this era of globalization, the delivery of information can be conveyed by any information media, this can also cause the spread of hoax information. This is very detrimental for recipients of information, especially information that is very important to receive and information related to the Covid or Corona virus that has hit the world and Indonesia. Based on data from the Gugus Tugas Percepatan Penanganan Covid-19, in the post-pandemic era 227 hoax information were still found that were identified. Therefore, in this study the authors conducted hoax information research, especially regarding the Covid-19 virus in the post-pandemic era by using the Text Mining method to find out information patterns that often appear in the post-Covid-19 pandemic. This paper will also show some results regarding the association of several words with the main word in hoax information in the post-Covid-19 pandemic era in the form of Term of Frequency, Word Cloud, Fruchterman Reingold Layout, and Circle Layout. The results of this study indicate that in the post-Covid-19 pandemic era, hoax information is frequently associated with ‘vaccines’ and ‘covid’, with 113 and 111 occurrences respectively. This is due to public skepticism regarding the safety and effectiveness of the Covid-19 vaccine, leading to the dissemination of hoaxes aimed at discouraging vaccination. Other commonly mentioned words in post-pandemic hoax information include ‘omicron’, ‘vaccinated’, ‘variant’, ‘pfizer’, ‘mRNA’, ‘virus’, and ‘vaccination’, albeit with lower frequencies.

Keywords: Covid-19; Text Mining; Hoax; Post Pandemic; Sentiment Analysis

1. PENDAHULUAN

Memasuki akhir tahun 2019, tepatnya pada bulan Desember, Covid-19 diidentifikasi pertama kali di Cina dan sejak itu menyebar di seluruh dunia hingga setidaknya sampai 187 negara dan beberapa wilayahnya [1]-[2]. Salah satu warga Wuhan, China dinyatakan positif terinfeksi virus Covid-19, penjual udang hidup di pasar ikan Kota Wuhan. Saat itulah dimulainya perkembangan virus tersebut sampai di berbagai Negara. Covid-19 adalah suatu virus yang dapat menyebar secara cepat. Beberapa hasil penelitian mengungkapkan cara penyebaran virus tersebut, salah satunya adalah World Health Organization (WHO) yang menerangkan bagaimana virus Covid-19 menyebar melalui tetesan kecil yang dikeluarkan dari hidung atau mulut saat orang yang terinfeksi bersin atau batuk [3].

Penyebaran virus COVID-19 ini bisa dikatakan sangat cepat, pada awal pandemi sampai tahun 2020 penduduk Indonesia yang terinfeksi Covid-19 telah mencapai 9.771 orang [4]. Berdasarkan penyebaran virus Covid-19 tersebut, akhirnya pemerintah beberapa negara membuat peraturan sebagai langkah pertama yang harus dilakukan yaitu berupa aturan social distancing. Akibat pandemi COVID-19 yang melanda dunia telah menimbulkan instabilitas di berbagai sektor, apalagi mengingat situasi saat ini yang memperlambat interkoneksi karena pada dasarnya dunia sedang mengalami “kejutan” atau belum siap menghadapi pandemi. Hal ini tentunya berdampak pada kepanikan dan kehati-hatian masyarakat dikarenakan ketidaktahuan masyarakat akan penanganan Covid-19 [5].

Pada awalnya beberapa negara maupun pemerintah tidak memberlangsungkan reaksi cepat tentang sosialisasi pencegahan seperti beberapa negara lainnya terkait dengan informasi yang didapatkan mengenai virus Covid-19. Hal ini dimungkinkan karena pencegahan rasa kekhawatiran setiap masyarakat pada negara tersebut, terkhususnya dengan berita hoaks yang menyebar dari beberapa orang yang tidak bertanggung jawab [6]. Memasuki era globalisasi, teknologi informasi menjadi salah satu media penyampaian data atau pengolahan informasi yang sangat cepat dan dapat menjangkau ke berbagai tempat, salah satu contohnya adalah media sosial

berbasis internet seperti Instagram, Facebook, WhatsApp, Line, KakaoTalk, dan Twitter. Hal ini memiliki manfaat yang positif dalam keperluan yang biasa kita lakukan untuk pemakaian yang sesuai dengan kegunaannya, akan tetapi media sosial ini juga bisa digunakan untuk menyebarkan berita hoaks yang belum ditentukan kebenarannya seperti permasalahan berita hoaks tentang virus Covid-19 yang penulis bahas saat ini. Melalui media sosial atau online mendapatkan, membagikan data atau informasi dari satu akun media sosial ke akun media sosial lainnya akan dengan mudahnya menjadi berita atau hal yang viral. Sebagai contoh, berita terkait Covid telah menguasai trending sepanjang hari. Beredar informasi dari internet dan media elektronik terkait penolakan warga terhadap penguburan jenazah penderita Covid-19. Padahal, tidak hanya pasien atau orang sakit dan keluarganya yang distigmatisasi, perawat yang merawat pasien Covid-19 juga mendapat stigmatisasi dari masyarakat. Bahkan ada banyak perawat yang diasingkan sampai diusir dari rumahnya, karena takut menjadi pembawa virus penyebab Covid-19 [7].

Hasil penelitian Juditha pada judul “Perilaku Masyarakat Terkait Penyebaran Hoaks Covid-19” menunjukkan bahwa teks merupakan jenis hoaks Covid-19 yang paling banyak dipercayai oleh responden, diikuti oleh gambar, video, dan audio. Lebih dari separuh responden (55%) mengatakan sulit membedakan apakah informasi yang mereka terima termasuk hoaks atau tidak, sementara 44% lainnya dapat mendeteksi berita palsu tersebut [8]. Hoaks Covid-19 yang paling banyak dipercayai meliputi isu pasien Covid-19 yang meninggal diisolasi, asal-usul virus, cara melindungi diri, cara mengobati virus, lokasi yang terpapar, serta cara mencegah penyebaran dan melindungi orang lain. Ada juga hoaks lainnya yang dipercayai oleh 12,30% responden [9]. Menanggapi penyebaran berita hoaks tersebut diperlukan kejelasan informasi yang disajikan sampai informasi diterima yang diterima oleh masyarakat dari pemerintah atau otoritas yang berwenang [10], akan tetapi hal tersebut juga belum tentu efektif dikarenakan penyebaran berita hoaks yang sangat cepat bahkan tidak terkendali sampai pada era pasca pandemi sekalipun [11].

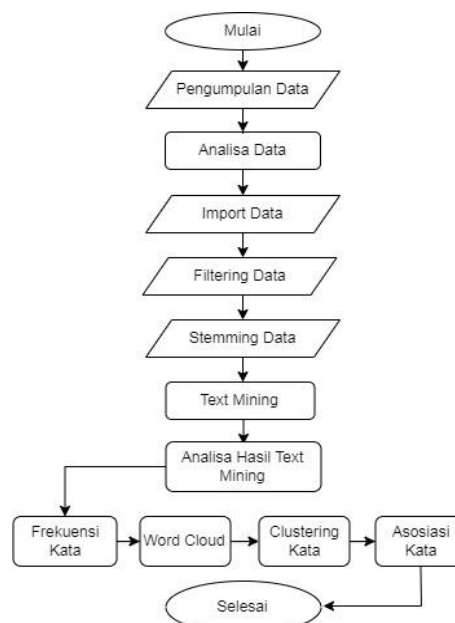
Metode yang dilakukan untuk menganalisis suatu Data Mining, salah satunya adalah Text Mining merupakan suatu proses ekstraksi informasi dari kumpulan data atau dokumen teks yang melibatkan penggunaan tools analisis sebagai bagian dari Data Mining [12], [13], [14]. Penelitian terdahulu telah berhasil melakukan deteksi hoaks pada media sosial dengan berbasis Text Mining dengan nilai akurasi sebesar 75% [15]. Penelitian selanjutnya juga telah melakukan klasifikasi topik yang diperlukan untuk pembaca berita online menggunakan Text Mining dengan nilai akurasi di atas 80% [16]. Oleh karena itu, metode Text Mining menjadi cara terbaik untuk mengetahui kata kunci penting pada informasi hoaks di era pasca pandemi.

Pada penelitian ini, penulis meneliti salah satu metode yang dapat membantu atau dilakukan untuk menganalisis atau mengeksplorasi setiap data, informasi maupun berita, dimana cukup ataupun sangat membantu dalam penanganan penyebaran berita hoaks khususnya tentang Covid-19 di era pasca pandemi wilayah Indonesia, yaitu dengan metode Text Mining, karena dalam penggunaan dan pemanfaatannya. Hal ini dapat dilakukan untuk penanganan berita hoaks ke depannya.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Berikut pada gambar 1 merupakan tahapan yang penulis lakukan dalam penelitian



Gambar 1. Alur Tahapan Penelitian

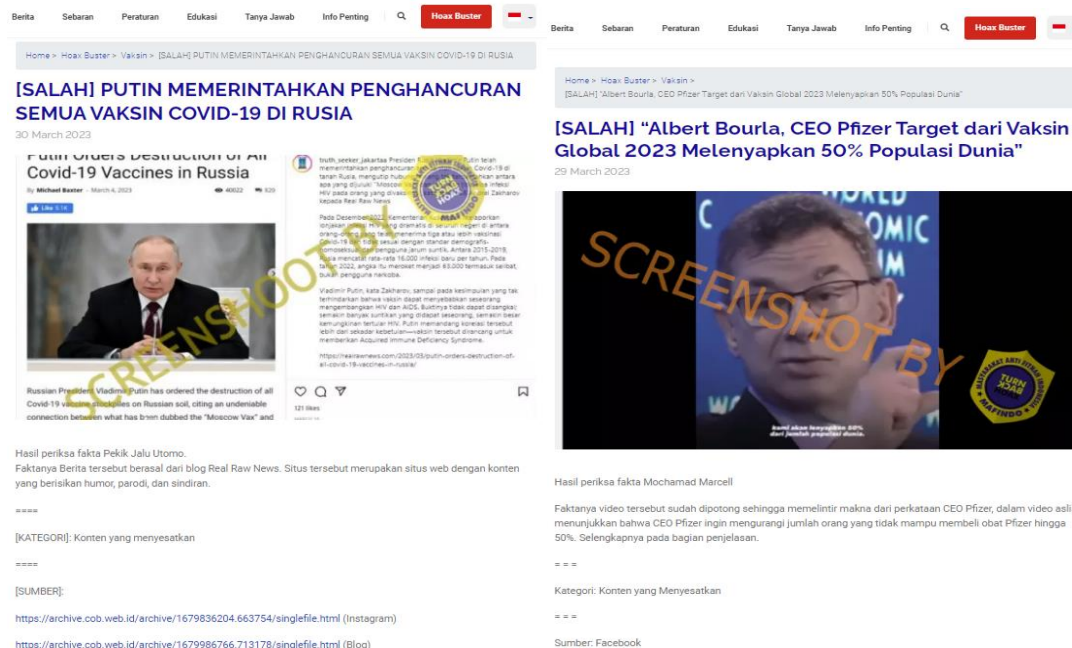
Dari gambar 1 proses analisis informasi hoaks menggunakan metode Text Mining, sebagai berikut:

- 1) Pengumpulan data yang diperoleh dari Gugus Tugas Percepatan Penanganan Covid-19 (covid19.go.id/hoax-buster) di era pasca pandemi (01 Januari 2022 – 31 Maret 2023),
- 2) Menganalisa data yang diperoleh untuk mengetahui karakteristik data informasi hoaks,
- 3) Melakukan import data, membuka file serta menampilkannya untuk diubah menjadi data dalam bentuk data frame atau seperti array. Kolom pertama (text) tersebut diubah menjadi corpus atau kumpulan text agar selanjutnya dapat diubah,
- 4) Melakukan proses Filtering Data, seperti: mengubah menjadi huruf kecil, menghapus tanda baca, menghapus angka, menghapus link URL, menambahkan kata henti, mengecualikan kata yang diperlukan (covid) dari kata henti,
- 5) Stemming Data, melakukan stemming text corpus, seperti menampilkan –me, -an menjadi tampil,
- 6) Analisa Frekuensi Kata, untuk menemukan kata yang memiliki frekuensi ≥ 10 agar selanjutnya bisa diplot menjadi barplot,
- 7) Wordcloud kata yang berukuran besar dan berada di tengah yang sering keluar di berita hoaks,
- 8) Clustering Kata menggunakan model Hierarchical Clustering, mengelompokkan kata berdasarkan keterhubungan setiap kata tersebut,
- 9) Menampilkan Asosiasi Kata menggunakan Layout Fruchterman Reingold, dan Layout Circle, menemukan kata-kata asosiasi seperti “omicron”, “pemerintah”, “pfizer”, dan “mrna” dalam informasi hoaks.

2.2 Materials

Pada penelitian kali ini, penulis menggunakan salah satu dataset resmi pemerintah yaitu, Gugus Tugas Percepatan Penanganan COVID-19 (covid19.go.id) yang menyediakan informasi data segala tentang Covid-19 di Indonesia maupun di dunia saat ini [4]. Website itu memiliki beberapa menu pilihan seperti, berita, sebaran, protokol, edukasi, tanya jawab, agenda, dan info-info lainnya serta mencakup peta sebaran, statistika perkembangan Covid-19 dalam berbagai pilihan aspek. Untuk informasi hoaks diperoleh pada bagian covid19.go.id/hoax-buster. Dari bagian tersebut terdapat ratusan informasi hoaks yang sudah terkonfirmasi kesalahannya mulai awal pandemi Covid-19 di negara Indonesia sampai dengan penulis mengumpulkan data (5 Mei 2023).

Informasi hoax diambil dengan menyalin teks informasi hoaks per judul hoaks mulai tanggal 01 Januari 2022 sampai dengan 31 Maret 2023. Data teks tersebut dipindah ke file excel sebagai data utama informasi hoaks pada era pasca pandemi. Total data yang telah terkumpul sebanyak 277 informasi hoax yang dimana setiap informasi hoax terdapat jenis kategori konten hoaksnya. Kategori informasi hoaks tersebut yaitu “Konteks yang Salah”, “Konten Palsu”, “Konten yang Menyesatkan”, dan “Satire”.



Gambar 2. Informasi Hoaks Era Pasca Pandemi

2.3 Informasi Hoaks

Hoax adalah berita palsu yang bertujuan menipu dan memperdaya pembaca/pendengarnya untuk mempercayainya. Berita palsu tersebut diciptakan oleh seseorang yang sadar bahwa berita tersebut tidak benar. Hoaks seringkali menggunakan istilah yang berbeda dengan barang/kejadian sebenarnya, dan bertujuan untuk mempengaruhi opini publik dan menguji kecerdasan dan ketelitian pengguna internet dan media sosial. Hoaks bisa disebar dengan

tujuan lelucon, black campaign, promosi palsu, ajakan untuk berbuat baik yang tidak berdasar, atau hanya untuk iseng. Namun, menyebarkan hoaks dapat menimbulkan ketidaktahuan yang luas jika banyak orang mempercayainya dan menyebarkan ke orang lain [17], [18].

Orang lebih mudah percaya hoaks jika informasinya mendukung keyakinannya. Misalnya, seseorang yang percaya bumi datar akan lebih mudah percaya pada teori konspirasi tentang foto satelit yang mendukung pandangannya. Orang cenderung merasa senang ketika opini atau keyakinannya mendapat dukungan, dan mungkin tidak terlalu peduli apakah informasi yang diterimanya valid atau tidak. Orang juga mudah menyebarkan informasi palsu jika kurang paham cara memeriksa kebenarannya di internet [19].

2.4 Text Mining

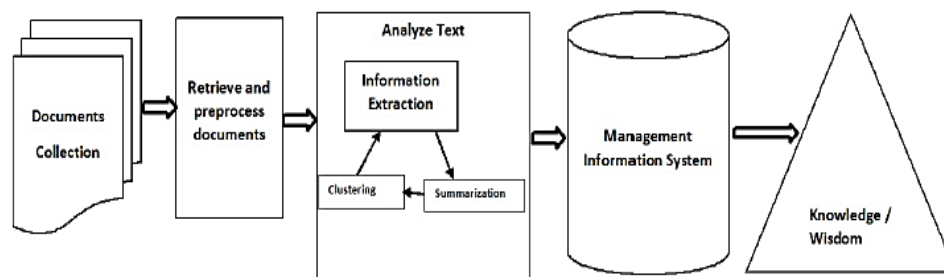
Untuk melakukan penelitian ini, berdasarkan penjelasan singkat pendahuluan di atas, penulis melakukan penanganan penyebaran berita hoaks tentang Covid-19 tersebut dengan menggunakan metode Text Mining. Metode yang dilakukan untuk menganalisis suatu Data Mining, salah satunya adalah Text Mining merupakan suatu proses ekstraksi informasi dari kumpulan data atau dokumen yang melibatkan penggunaan tools analisis sebagai bagian dari Data Mining. Salah satu fungsi utama dari Text Mining adalah melakukan kategorisasi pada data atau dokumen tersebut [20]. Perbedaan Data Mining dan Text Mining adalah proses penggunaan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengeksplorasi dan menganalisis data dalam database besar dengan tujuan mengidentifikasi pola, hubungan, dan informasi yang bermanfaat. Proses ini melibatkan pengekstrakan pengetahuan yang tersembunyi dalam data yang tidak terstruktur atau terstruktur.

Tahapan Text Mining meliputi, Text Preprocessing dan Feature Selection. Berikut penjelasan tahap-tahap tersebut:

- 1) **Text preprocessing** merujuk pada tahap awal dari proses Text Mining yang melibatkan serangkaian rutinitas dan persiapan data yang akan digunakan dalam operasi knowledge discovery sistem Text Mining. Tahap ini termasuk beberapa tindakan seperti toLowerCase, yaitu mengonversi semua karakter huruf menjadi huruf kecil, dan tokenizing, yaitu proses memisahkan deskripsi yang terdiri dari kalimat-kalimat menjadi kata-kata individual. Selain itu, tahap ini juga melibatkan penghapusan delimiter seperti tanda titik (.), koma (,), spasi, dan karakter angka pada kata-kata tersebut.
- 2) **Feature Selection** merupakan suatu teknik yang bertujuan untuk mengurangi dimensi dari sebuah kumpulan teks atau dokumen dengan cara menghapus kata-kata yang dianggap tidak penting atau tidak relevan dengan isi dokumen tersebut. Hal ini dilakukan agar proses klasifikasi teks menjadi lebih efektif dan akurat. Pada tahap Feature Selection, dilakukan beberapa proses seperti penghapusan stopword dan stemming terhadap kata-kata yang berlebihan. Proses stopword removal dilakukan dengan cara menghilangkan kata-kata yang sering muncul namun tidak memiliki makna atau tidak relevan dengan konteks dokumen. Sedangkan proses stemming dilakukan dengan cara mengubah kata-kata yang memiliki akar kata yang sama menjadi bentuk dasarnya.

2.5 Proses Text Mining

Pada Text Mining terdapat beberapa proses umum yang harus dilakukan seperti yang di gambarkan pada gambar 3. Berikut penjelasan proses umum tersebut:



Gambar 3. Proses Text Mining

Dari gambar 3 dapat dijelaskan sebagai berikut:

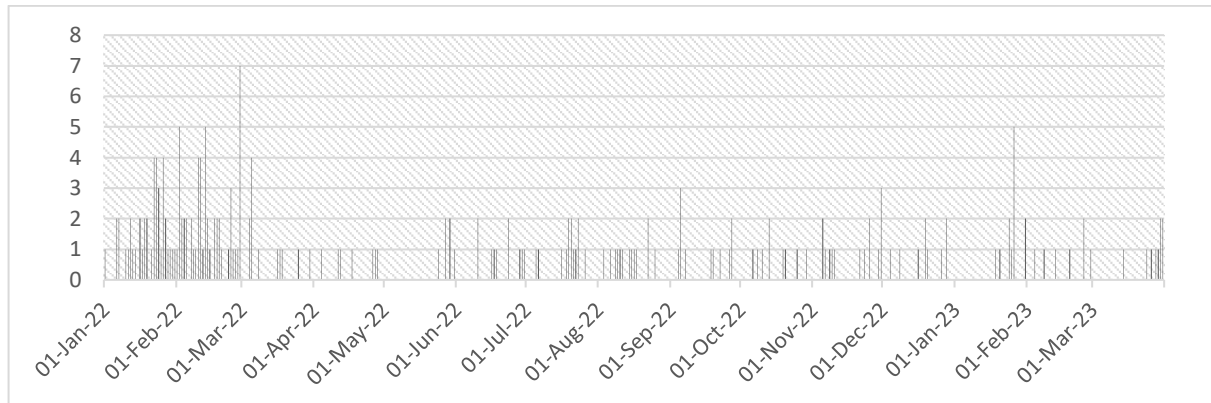
- 1) **Documents Collection:** Pengumpulan data yang tidak terstruktur berasal berbagai sumber yang ditentukan baik dalam bentuk format file seperti teks biasa, halaman web, file .pdf, dan lain-lain,
- 2) **Retrieve and preprocess documents:** Operasi pra-pemrosesan dan pembersihan yang dilakukan untuk mendeteksi dan menghilangkan anomali. Memastikan mendapatkan esensi nyata dari teks yang tersedia dan melakukan proses mengidentifikasi akar kata tertentu,
- 3) **Analyze Text:** Pemrosesan dan pengendalian algoritma yang diterapkan untuk mengaudit dan memproses otomatis kumpulan data sampai dapat digunakan sistem,
- 4) **Management Information System:** Menganalisis pola yang diimplementasikan oleh Management Information System (MIS) atau Sistem Informasi Manajemen,

- 5) **Knowledge/Wisdom:** Informasi yang diproses berdasarkan tahapan sebelumnya digunakan untuk mengekstraksi informasi yang diperlukan secara relevan untuk melakukan penentuan keputusan.

3. HASIL DAN PEMBAHASAN

3.1 Analisa Data

Dari data yang terkumpul terdapat 227 informasi hoaks di era pasca pandemi (01 Januari 2022 – 31 Maret 2023). Data tersebut memiliki sebaran jumlah informasi hoaks yang beragam setiap harinya yang dapat dilihat pada Gambar 4 berikut:



Gambar 4. Frekuensi Informasi Hoaks Covid-19 Per-hari

Terlihat pada Gambar 4, frekuensi informasi hoaks pada antara Bulan Januari 2022 sampai Maret 2022 memiliki jumlah informasi hoaks yang cukup banyak daripada di tanggal-tanggal lainnya. Namun demikian, informasi hoaks tentang Covid-19 tidak pernah surut sampai dengan 31 Maret 2023.

Data informasi hoaks terdapat empat kategori yaitu “Konteks yang Salah”, “Konten Palsu”, “Konten yang Menyesatkan”, dan “Satire”. Jumlah informasi kategori “Konteks yang Salah” sebanyak 22 informasi yang dapat dilihat contohnya pada tabel 1 berikut:

Tabel 1. Contoh Informasi Hoax Kategori “Konteks yang Salah”

Kategori	Text Informasi Hoaks	Tanggal
Konteks yang salah	Penduduk Aborigin di Australia Menyambut Pemerintah dengan Busur dan Panah sebagai Bentuk Penolakan Vaksin Covid-19	07-Jan-22
Konteks yang salah	Protes Melawan Tirani Covid-19 di Belanda	22-Jan-22
Konteks yang salah	“#StopVaccineForTheKids Apa urgensinya?”	26-Jan-22
Konteks yang salah	“HARI RAYA IMLEK, HADIR & BERKERUMUN TANPA PROKES sdh diCONTOHkan oleh Jokowi”	02-Feb-22
Konteks yang salah	Dzalim, Cara Vaksin Model Baru Anti Omicron	07-Feb-22
Konteks yang salah	Lampion untuk Perayaan Imlek Saat Pandemi Covid-19	07-Feb-22

Konteks yang salah terjadi ketika informasi atau data diambil dari suatu konteks yang salah atau tidak sesuai dengan kenyataan. Misalnya, kutipan atau gambar yang diambil dari konteks yang berbeda, sehingga memberikan kesan yang salah atau tidak akurat tentang suatu topik atau kejadian. Kategori “Konteks yang Salah” biasanya digunakan dalam pengecekan fakta untuk memeriksa kebenaran sebuah informasi atau klaim yang tersebar di media sosial atau platform online lainnya.

Sedangkan banyaknya informasi kategori “Konten Palsu” sebanyak 18 informasi yang dapat dilihat contohnya pada tabel 2 berikut:

Tabel 2. Contoh Informasi Hoax Kategori “Konten Palsu”

Kategori	Text Informasi Hoaks	Tanggal
Konten Palsu	Penerima BLT UMKM Bulan Januari 2022 Sebesar Rp1,2 Juta	06-Jan-22
Konten Palsu	Lowongan Kerja Vaksinator di Puskesmas Kota Batu	14-Jan-22

Kategori	Text Informasi Hoaks	Tanggal
Konten Palsu	Jasa Pembuatan Sertifikat Vaksin Tanpa Harus Melakukan Vaksinasi	23-Jan-22
Konten Palsu	Bantuan Sosial Kemensos untuk Dosen, Guru, Siswa, dan Mahasiswa	01-Feb-22
Konten Palsu	Razia Masker dengan Denda Rp250 Ribu di Seluruh Indonesia	04-Feb-22
Konten Palsu	Sri Sultan Hamengku Buwono X Mengganggu Covid 19 Sandiwara di Bumi NKRI	05-Feb-22

Konten palsu digunakan untuk menggambarkan informasi atau berita yang tidak benar atau tidak akurat. Konten Palsu ini bisa terdiri dari berbagai jenis, termasuk hoaks, berita palsu, atau informasi yang keliru. Konten Palsu biasanya disebarkan dengan tujuan tertentu, seperti menyebarkan opini atau pandangan tertentu, menimbulkan kepanikan, atau merusak citra suatu pihak atau organisasi.

Selanjutnya data informasi hoaks didominasi oleh kategori “Konten yang Menyesatkan” yaitu sebanyak 183 informasi yang dapat dilihat contohnya pada Tabel berikut:

Tabel 3. Contoh Informasi Hoax Kategori “Konten yang Menyesatkan”

Kategori	Text Informasi Hoaks	Tanggal
Konten yang Menyesatkan	Gagal Jantung hingga Strok adalah Gejala Omicron	01-Jan-22
Konten yang Menyesatkan	Vaksin mRNA pada Pfizer dan Moderna Berbahaya Bagi Anak-Anak	06-Jan-22
Konten yang Menyesatkan	Warga Korea Selatan yang Belum Vaksin Dilarang Mengikuti Pemilihan Presiden	07-Jan-22
Konten yang Menyesatkan	CDC Mengumumkan Tes PCR Tidak Mampu Membedakan Virus SARS-Cov-2 dan Influenza	10-Jan-22
Konten yang Menyesatkan	WHO Mengatakan Vaksinasi di Seluruh Dunia Harus Dihentikan	11-Jan-22
Konten yang Menyesatkan	Benda Logam Menempel pada Lengan setelah divaksin	12-Jan-22

Konten yang menyesatkan bisa berupa berita palsu, klaim tidak terbukti, manipulasi gambar atau video, atau tafsiran yang keliru dari data atau fakta. Konten yang menyesatkan ini sering kali disebarkan dengan tujuan tertentu, seperti untuk mempromosikan pandangan atau opini tertentu, menimbulkan kepanikan, atau mempengaruhi pandangan orang terhadap suatu topik atau kejadian. Terakhir, kategori informasi hoax pada data adalah Satire dengan sebanyak 4 informasi yang dapat dilihat contohnya pada tabel 4 berikut:

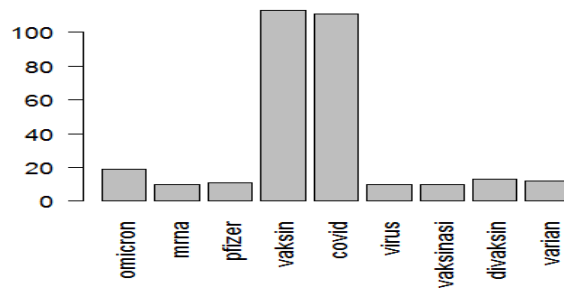
Tabel 4. Contoh Informasi Hoax Kategori “Satire”

Kategori	Text Informasi Hoaks	Tanggal
Satire	Kipas Angin Listrik yang Diletakkan di Dahi Dapat Menghilangkan Virus	22-Jan-22
Satire	Bill Gates akan Menarik Semua Peredaran Vaksin Covid-19	23-Jan-22
Satire	Justin Bieber Mengatakan Kelumpuhan Wajahnya Disebabkan Oleh Vaksin Covid-19	23-Jun-22
Satire	Pembaruan Aplikasi Peduli Lindungi	08-Feb-23

Satire adalah jenis tulisan atau karya seni yang menggunakan humor, ironi, dan lelucon untuk mengekspos kelemahan atau masalah dalam masyarakat, politik, atau budaya. Kategori “Satire” seringkali menggunakan bahasa dan narasi yang berlebihan atau berlawanan, sehingga terlihat sangat jelas bahwa karya tersebut tidak dimaksudkan untuk diambil secara harfiah. Namun, terkadang kategori “Satire” dapat diterima atau dipahami dengan cara yang salah oleh sebagian orang, terutama jika tidak disertai dengan penjelasan yang jelas atau dikaitkan dengan konteks yang tepat.

3.2 Analisa Frekuensi Kata (Term of Frequency)

Berdasarkan gambar 5 grafik frekuensi (diagram batang) di atas dapat diketahui frekuensi dari kata yang sering muncul bersamaan dengan topik informasi hoaks di era pandemi Covid-19. Dimana frekuensi kata yang sering muncul akan digambarkan dengan tingginya suatu grafik batang tersebut. Kata yang paling banyak pada informasi hoaks ini adalah kata “vaksin” sebanyak 113 kata dan “covid” sebanyak 111 kata. Dari hasil tersebut sudah dapat disimpulkan bahwa informasi hoaks sebagian besar berhubungan dengan vaksin Covid-19. Hal ini dikarenakan masyarakat yang belum yakin akan vaksin Covid-19 sehingga memunculkan tindakan untuk menghindari diri dengan memberikan informasi hoaks.



Gambar 5. Barchart Frekuensi Kata Informasi Hoax

Sedangkan kata-kata lain yang sering muncul pada informasi hoaks di era pasca pandemi adalah kata “omicron” sebanyak 19 kata, “divaksin” sebanyak 13 kata, “varian” sebanyak 12 kata, “pfizer” sebanyak 11 kata, “mrna” sebanyak 10 kata, “virus” sebanyak 10 kata, dan “vaksinasi” sebanyak 10 kata. Kata-kata tersebut merupakan topik-topik menarik dan berhubungan dengan era pandemi Covid-19 telah dibahas juga oleh masyarakat.

3.3 Word Cloud

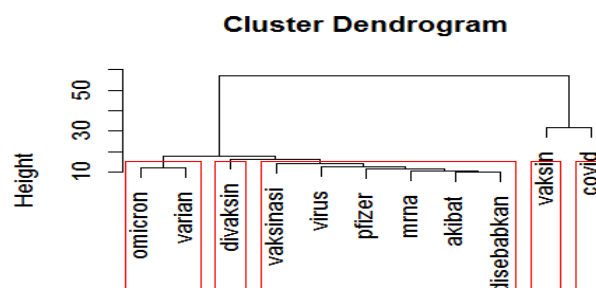
Gambar 6 di bawah menunjukkan suatu Word Cloud yang berkaitan dengan informasi hoaks di pasca pandemi Covid-19. Sama hal nya seperti Analisa Frekuensi Kata, dimana menunjukkan kata-kata atau topik-topik yang muncul berkaitan langsung dengan topik hoaks era pasca pandemi. Word Cloud menampilkan kata-kata yang terkumpul dengan bentuk seperti awan. Banyaknya jumlah kata yang muncul, dapat diketahui dengan ukuran besar huruf yang ditampilkan dari word cloud, semakin besar ukuran huruf pada menandakan bahwa kata tersebut adalah kata yang sering muncul pada informasi hoaks di era pasca pandemi.



Gambar 6. Word Cloud Informasi Hoax Era Pasca Pandemi Covid-19

Terdapat kata-kata lain yang muncul pada Word Cloud selain barchart Frekuensi Kata pada Gambar 4. Isu-isu yang sering muncul juga ada pada kata “modern” sebagai salah satu vaksin Covid-19. Kemudian “Cacar Monyet” yang dikambing hitamkan sebagai efek dari vaksin Covid-19. Kata “pemerintah” pun dijadikan isu hangat yang disebar oleh penyebar informasi hoaks untuk menyakinkan informasi hoaksnya.

3.4 Clustering Kata



Gambar 7. Clustering Kata Informasi Hoax Era Pasca Pandemi Covid-19

Kata-kata informasi hoaks dikelompokkan dengan model Hierarchical Clustering yang menunjukkan keterikatan kata pada setiap kelompoknya. Pada Gambar di atas merupakan hasil clustering kata yang membagi kelompok menjadi 5 kelompok kata. Daftar kata pada setiap kelompok tersebut yaitu sebagai berikut:

1) **Kelompok 1** : kata “varian omicron”

Kelompok ini memperlihatkan kata “varian omicron” sering muncul pada informasi hoaks dikarenakan varian tersebut merupakan varian Covid-19 yang ada pada era pasca pandemi,

2) **Kelompok 2** : kata “divaksin”

Kelompok ini menunjukkan isu yang sering dimunculkan untuk menghindari vaksinasi maka berita dengan kata “divaksin” selalu dimanipulasi pada informasi hoaks,

3) **Kelompok 3** : kata “vaksinasi”, “virus”, “pfizer”, “mrna”, “akibat”, “disebabkan”

Kelompok ini lebih menunjukkan seringnya manipulasi informasi tentang vaksin jenis pfizer yang merupakan vaksin booster yang diwajibkan oleh pemerintah,

4) **Kelompok 4** : kata “vaksin”

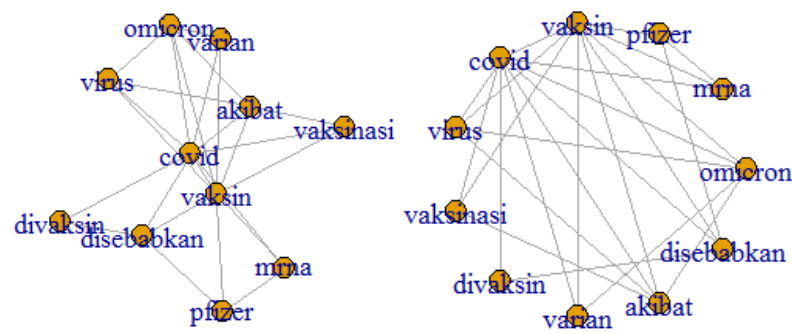
Sama halnya dengan kelompok 2, kata “vaksin” sering berkaitan dengan kata-kata pada informasi hoaks era pasca pandemi dikarenakan ketakutan masyarakat akan vaksinasi,

5) **Kelompok 4** : kata “covid”

Kelompok ini merupakan objek utama pada informasi hoaks yaitu nama virus itu sendiri.

3.5 Asosiasi Term

Selain pengelompokkan kata dalam bentuk Cluster Dendrogram, relasi setiap kata dapat juga dilihat hasilnya dalam bentuk beberapa Layout Fruchterman Reingold dan Layout Circle, sebagai berikut:



Gambar 8. Visualisasi Relasi Setiap Kata Informasi Hoaks dengan Layout Fruchterman Reingold (bagian kiri) dan Layout Circle (bagian kanan)

Gambar 8 memvisualisasikan relasi antar kata yang mengandung informasi hoaks era pasca pandemi menggunakan tata letak fruchterman-reingold dan Layout Circle. Node kata pada layout fruchterman-reingold yang berada di tengah gambar menunjukkan kata yang paling sering berkaitan dengan informasi hoaks era pasca pandemi seperti kata “covid” dan “vaksin”.

Setelah mengetahui kata yang sering muncul pada informasi hoaks di era pasca pandemi yaitu “pfizer”, “mrna”, “omicron”, dan “pemerintah”. Maka perlu dianalisa lebih dalam lagi kata-kata apa yang berasosiasi dengan empat kata tersebut. Berikut adalah daftar kata yang berasosiasi dengan kata “pfizer” dan kata “mrna” beserta persentase asosiasinya. Nilai persentase asosiasi menunjukkan seberapa keterikatan kata-kata tersebut, semakin besar nilai persentase asosiasi maka semakin besar keterikatan.

Tabel 5. Term yang Berasosiasi dengan Term “pfizer” dan “mrna”

Term Yang Berasosiasi dengan Term 'pfizer'	Persentase Asosiasi	Term Yang Berasosiasi dengan Term 'mrna'	Persentase Asosiasi
ceo	0.44	hewan	0.58
efektivitas	0.42	memakan	0.58
klorida	0.29	ternak	0.58
membahayakan	0.29	tervaksinasi	0.58
potasium	0.29	lansia	0.32
terbaru	0.29	kehamilan	0.28
beredar	0.29	dna	0.28
mendistribusikan	0.29	makanan	0.28

Kata-kata yang berkaitan erat dengan kata “pfizer” lebih banyak menunjukkan opini palsu tentang efektivitas vaksin Pfizer, berbahaya vaksin Pfizer, serta manajemen produksi Pfizer. Kemudian kata-kata yang berkaitan erat dengan kata “mrna” lebih banyak menunjukkan informasi palsu tentang gejala palsu dari cara kerja vaksin jenis Pfizer dan Moderna yang berbasis messenger RNA (mRNA).

Tabel 6. Term yang Berasosiasi dengan Term “omicron” dan “pemerintah”

Term Yang Berasosiasi dengan Term 'omicron'	Persentase Asosiasi	Term Yang Berasosiasi dengan Term 'pemerintah'	Persentase Asosiasi
varian	0.50	aborigin	0.44
jahe	0.38	busur panah	0.44
kopi	0.31	penolakan	0.44
mengonsumsi	0.31	euthanasia	0.44
pahit	0.31	new zealand	0.44
sembuhkan	0.31	antena	0.44
sinar	0.31	peru	0.44
uv	0.31	haji	0.44
chemtrail	0.31	baju	0.44
keracunan	0.31	hazmat	0.44
campuran	0.31	kebijakan	0.44
daun	0.31	tiongkok	0.44
pepaya	0.31	penduduk	0.31
paracetamol	0.31	dana	0.31

Kata-kata yang berkaitan erat dengan kata “omicron” lebih banyak menunjukkan informasi palsu untuk menangani Covid-19 varian Omicron dengan memberikan opini bahwa varian tersebut dapat disembuhkan dengan Paracetamol, Sinar UV, Daun Pepaya, Jahe, dan Kopi. Kemudian kata-kata yang berkaitan erat dengan kata “pemerintah” lebih banyak menunjukkan informasi palsu tentang penanganan Covid-19 pada pemerintah di wilayahnya. Wilayah-wilayah yang terkait pada informasi hoaks di era pasca pandemi adalah Aborigin, New Zealand, Peru, dan Tiongkok.

4. KESIMPULAN

Berdasarkan hasil penelitian, analisis, dan pengujian yang dilakukan terhadap penerapan metode Text Mining pada informasi hoaks di era pasca pandemi Covid-19, ditemukan bahwa terdapat keterkaitan antara hoaks di era pasca pandemi dengan kata kunci “vaksinasi” dan “covid”. Melalui tahapan Text Preprocessing, kata-kata yang tidak relevan dihapus menggunakan tahapan Feature Selection. Dokumen-dokumen hoaks yang diperoleh dari website Gugus Tugas Percepatan Penanganan COVID-19 kemudian dianalisis dan diproses untuk menghasilkan kata-kata penting. Pada era pasca pandemi Covid-19, kata-kata yang paling sering muncul dalam informasi hoaks adalah “vaksin” dan “covid” dengan jumlah masing-masing 113 kata dan 111 kata. Informasi hoaks yang tersebar terutama berkaitan dengan vaksin Covid-19 karena adanya keraguan masyarakat terhadap keamanan dan efektivitas vaksin. Selain itu, kata-kata lain yang sering muncul dalam informasi hoaks di era pasca pandemi adalah “omicron” (19 kata), “divaksin” (13 kata), “varian” (12 kata), “pfizer” (11 kata), “mrna” (10 kata), “virus” (10 kata), dan “vaksinasi” (10 kata). Semua topik ini menarik perhatian masyarakat karena keterkaitannya yang erat dengan pandemi Covid-19 dan telah menjadi perbincangan luas. Oleh karena itu, penting untuk selalu memeriksa kebenaran informasi sebelum menyebarkannya agar tidak memperburuk situasi pandemi yang sedang berlangsung.

REFERENCES

- [1] L. F. Moriarty, “Public Health Responses to COVID-19 Outbreaks on Cruise Ships Worldwide February March 2020”, US Department of Health and Human Services/Centers for Disease Control and Prevention MMWR, 69(12), pp. 347-352, 2020.
- [2] H. Shi, “Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study”, The Lancet Infectious Diseases. Elsevier Ltd, 20(4), pp. 425–434, 2020, doi: 10.1016/S1473-3099(20)30086-4.
- [3] C. K. L. Pastor, “Sentiment Analysis on Synchronous Online Delivery of Instruction due to Extreme Community Quarantine in the Philippines caused by COVID-19 Pandemic”, Asian Journal of Multidisciplinary Studies, 3(1), 2020.
- [4] G. T. COVID-19, “Gugus Tugas Percepatan Penanganan COVID-19”, 2023, [Online]. Retrieved April 01, 2023, from Gugus Tugas Percepatan Penanganan COVID-19: <https://Covid19.go.id/>.
- [5] N. R. Yunus and A. Rezki, “Kebijakan Pemberlakuan Lock Down Sebagai Antisipasi Penyebaran Corona Virus Covid-19”, Salam: Jurnal Sosial & Budaya Syar-i, 7(3), pp. 227-238, 2020.
- [6] N. H. Zahro, G. A. Setiawan, and A. Amroellah, “POLA HIDUP SEHAT PASCA PANDEMI COVID-19”, MIMBAR INTEGRITAS: Jurnal Pengabdian, 1(2), pp. 299-306, 2022.
- [7] R. Abudi, Y. Mokodompis, and A. N. Magulili, “Stigma Terhadap Orang Positif COVID-19”, Jambura Journal of Health Sciences and Research, 2(2), pp. 77-84, 2020.
- [8] C. Juditha, “Perilaku Masyarakat Terkait Penyebaran Hoaks Covid-19”, Jurnal Pekommas, 5(2), 2020.
- [9] C. Juditha, “Interaksi Komunikasi Hoax di Media Sosial serta Antisipasinya”, Jurnal Pekommas, 3(1), 2018.
- [10] M. F. Soleh, “Penimbunan Alat Pelindung Diri pada Masa Pandemi Covid-19: Kajian Hukum Pidana Bidang Perlindungan Konsumen”, Undang: Jurnal Hukum, 3(1), 2020, doi: 10.22437/ujh.3.1.1-31.
- [11] K. H. Simatupan, B. N. E. Motos, and T. P. Kelung, “Analisis Nasib Globalisasi Pasca Pandemi Covid-19”, UPH Journal of International Relations, pp. 30-37, 2022.

- [12] Y. Pernando, E. L. Febrianti, I. Syafrinal, Y. Roza, and U. F. Afifah, “Deep Learning for Faces on Orphanage Children Face Detection”, *Jurteksi*, vol. 9, no. 1, pp. 25-32, 2022.
- [13] A. R. Purnajaya, V. Lieputra, V. Tayanto, and J. G. Salim, “Implementasi Text Mining untuk Mengetahui Opini Masyarakat Tentang Climate Change”, *Journal of Information System and Technology*, vol. 3, no. 3, pp. 36-44, 2022.
- [14] A. R. Purnajaya, V. O. Noverio, and C. Alvaro, “Analisis Ancaman COVID-19 Varian XBB di Indonesia pada Jejaring Media Sosial Twitter menggunakan Text Mining”, *JITTER*, vol. 4, no. 1, pp. 1567-1575, 2023.
- [15] A. G. Tammam, “Deteksi Hoaks pada Media Sosial Berbasis Text Mining Classification System”, *Sistem Informatika - Universitas Nusantara PGRI Kediri*, 2018.
- [16] O. Sihombing, S. T. Y. Sitorus, E. Indra, and S. H. Sinurat, “Comparative Analysis of Indonesian Text Mining News Online Classification Using the K-Nearest Neighbor and Random Forest Algorithm”, *JUSIKOM Prima*, vol. 6, no. 1, 2022.
- [17] A. Ishak, Y. Y. Chen, and S. Yong, “Distance-based Hoax Detection System”, 2012 International Conference on Computer & Information Science (ICCIS), Kuala Lumpur, Malaysia, pp. 215-220, 2012, doi: 10.1109/ICCISci.2012.6297242.
- [18] D. R. Rahadi, “Perilaku Pengguna dan Informasi Hoax di Media Sosial”, *Jurnal Manajemen & Kewirausahaan*, 5(1), pp. 58-70, 2017.
- [19] A. Rahmadhany, A. A. Safitri, and D. Irwansyah, “Fenomena Penyebaran Hoax dan Hate Speech pada Media Sosial”, *Jurnal Teknologi dan Informasi Bisnis*, 3(1), pp. 30-43, 2021, doi: 10.47233/jteksis.v3i1.182.
- [20] I. Adiyana and F. Hakim, “Implementasi Text Mining Pada Mesin Pencarian Twitter Untuk Menganalisis Topik-Topik Terkait Kpk Dan Jokowi”, *Prosiding Seminar Nasional Matematika dan Pendidikan Matematika UMS 2015*, pp. 570-581, 2015.