

## PENERAPAN METODE RANDOM FOREST DALAM MENDETEKSI BERITA HOAX

Tio Tambunan, Margaretha Yohanna✉, Arina Prima Silalahi

Program Studi Teknik Informatika, Universitas Methodist Indonesia, Medan, Indonesia

Email: [yohanna.na2@gmail.com](mailto:yohanna.na2@gmail.com)

DOI: <https://doi.org/10.46880/jmika.Vol7No2.pp301-306>

### ABSTRACT

*Hoax is information that is not true. The Ministry of Communication and Informatics (Kominfo) found that there was 2,099 hoax news that was spread thousands of times via social media. This generally impacts the community so it can lead to a crisis of confidence in the government. This arises because many message recipients have different literacy levels, which will affect how people analyze the information conveyed. This research uses the Random Forest method, which is used to classify large amounts of data to detect hoax news. The research results show that the Random Forest method is proven to be able to classify hoax news based on data that has been weighted and entered into the system. From the results of the study using 200 data sets, which were divided by 80% in the form of training data and 20% of testing data, the classification results obtained from the testing data were in the form of 28 positive sentiments and 23 negative sentiments with an accuracy rate of 98%.*

**Keyword:** Hoax, Classification, Random Forest.

### ABSTRAK

*Hoax adalah informasi yang tidak benar. Kementerian Komunikasi dan Informatika (Kominfo) menemukan ada 2.099 berita hoax yang disebarakan ribuan kali melalui media sosial. Hal tersebut secara umum berdampak pada masyarakat sehingga dapat menimbulkan krisis kepercayaan terhadap pemerintah. Hal tersebut muncul karena banyak penerima pesan memiliki tingkat literasi yang berbeda sehingga akan mempengaruhi bagaimana masyarakat dalam menelaah informasi yang disampaikan. Metode Random Forest dalam penelitian ini digunakan untuk mengklasifikasikan data dalam jumlah besar untuk mendeteksi berita hoax. Hasil penelitian menunjukkan bahwa metode Random Forest terbukti dapat mengklasifikasikan berita hoax berdasarkan data yang diinputkan kedalam sistem. Penelitian ini menggunakan 200 data set, dimana dibagi menjadi 80% berupa data training dan 20% berupa data testing. Hasil klasifikasi dari data testing didapatkan sentiment positif sebanyak 28 dan sentimen negative sebanyak 23 dengan tingkat akurasi sebesar 98%.*

**Kata Kunci:** Hoax, Klasifikasi, Random Forest.

### PENDAHULUAN

Hoax merupakan berita atau informasi yang disampaikan dengan benar (Susanti et al., 2022). Berdasarkan cambridge dictionary, kata hoax berarti tipuan atau lelucon. Kegiatan untuk menipu, melakukan trik penipuan maupun rencana penipuan sering dikenal sebagai hoax. Saat ini, masyarakat yang sudah tidak asing lagi dengan penggunaan media sosial. Media sosial banyak digunakan untuk mencari informasi apa saja yang berlangsung saat ini (Bahri, 2021). Saat ini, isu mengenai hoax Covid-19 masih banyak mendominasi berbagai pemberitaan di media sosial (Edy & Florina, 2021). Kementerian Komunikasi dan Informatika menemukan sekitar 2.099 berita terkait Covid-19 yang telah disebarakan sebanyak ribuan kali dalam jangka waktu dari 23 Januari 2020 – 6 Februari 2022. Dari ratusan kasus penyebaran berita hoax yang

masuk ke ranah hukum, yang disebarakan melalui sosial media Facebook, ada sebanyak 4.630 berita yang sudah diblokir dan 151 berita sebaran sedang ditindak lanjuti.

Media sosial merupakan salah satu wadah yang sangat rentan dan sering disalahgunakan oleh oknum yang tidak bertanggung jawab untuk menyebarkan sebuah berita hoax (Manalu et al., 2022). Menurut data yang sudah dikumpulkan, penyebaran berita hoax paling banyak disebarakan melalui media sosial dan Facebook menempati urutan tertinggi yaitu sebesar 92,40%, aplikasi chatting sebesar 62,80%, dan situs web lainnya sebesar 34,90%. Berdasarkan hasil riset yang dilakukan oleh We Are Social, menyebutkan bahwa pengguna Facebook di Indonesia mencapai 122 juta dan pengguna Twitter mencapai 78 juta terhitung per November 2019. Terkait banyaknya berita hoax yang disebarakan melalui media sosial, bila hal tersebut

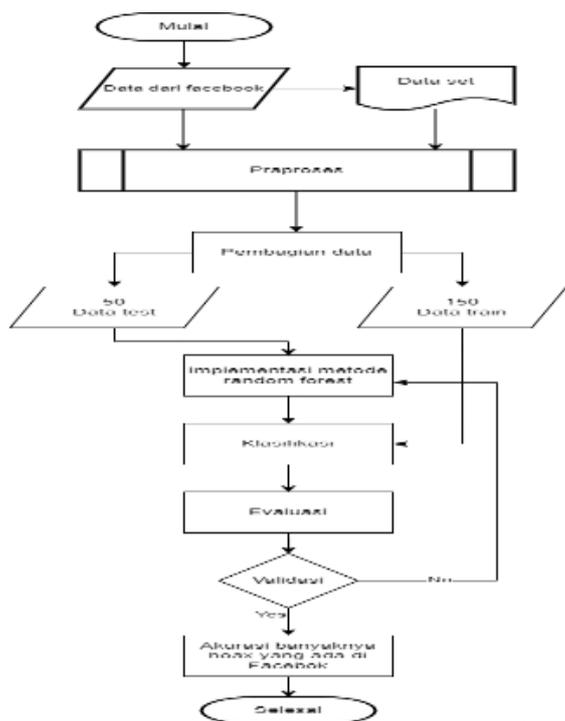
dibiarkan secara terus menerus, akan menimbulkan dampak pada masyarakat secara umum dimana masyarakat akan mengalami krisis kepercayaan terhadap Pemerintah karena banyaknya berita hoax tersebut. Krisis kepercayaan dapat timbul karena didalam proses komunikasi, proses penerimaan pesan yang berbeda, dengan tingkat literasi yang berbeda dapat mempengaruhi masyarakat dalam berpikir dan bertindak.

Hal yang terpenting dari penelitian ini adalah seringkali informasi yang disampaikan melalui media sosial merupakan informasi yang tidak akurat, sehingga untuk mendapatkan informasi yang lebih akurat, harus diperoleh melalui saluran resmi yang dikeluarkan oleh pemerintah. Untuk mengatasi hal tersebut, diperlukan sebuah alat bantu yang dapat digunakan untuk mendeteksi berita hoax secara otomatis, sehingga nantinya masyarakat dapat lebih berhati-hati dalam membaca dan mencerna berita yang disampaikan melalui media sosial (Roflin et al., 2021). Penelitian ini bertujuan untuk mengetahui apakah berita yang disebarkan melalui media sosial merupakan berita hoax atau tidak, sehingga nantinya masyarakat tidak mudah terprovokasi.

**METODE PENELITIAN**

**Framework penelitian**

Tahapan ini terdiri dari beberapa tahapan, adapun tahapan-tahapan yang akan dilakukan dapat dilihat pada Gambar 1.



Gambar 1. Framework Penelitian

**Pengumpulan Data**

Pada tahap ini dilakukan pengumpulan data mengenai seberapa besar kemungkinan masyarakat dapat percaya dengan adanya hoax yang tersebar di media sosial dan seberapa banyak orang yang sudah terpengaruh dengan berita hoax tersebut. Adapun pengumpulan data pada penelitian ini yaitu:

**Pengumpulan Data Hoax**

Data berita hoax diambil dari website Kominfo. Data hoax tersebut dapat dilihat pada Tabel 1.

Tabel 1. Data Hoax Kominfo

No.	Test
1	Beredar sebuah unggahan video ... terkait dengan perebutan Pulau Pasir.
2	Beredar unggahan di media sosial Twitter ... di Cianjur, pada Senin, 21 November 2022.
3	Telah beredar pesan berantai WhatsApp ... Piala Dunia Qatar.
4	Beredar di media sosial, sebuah foto ... Rumah Sakit Umum Daerah (RSUD) Sayang Cianjur
5	Beredar sebuah video di media sosial ... Amerika Serikat usai mengikuti KTT G20.
6	Beredar di media sosial Facebook ... sinyal 5G untuk mengumpulkan dan mengirimkan data.

**Pengumpulan Data Validasi Pada Media Mainstream**

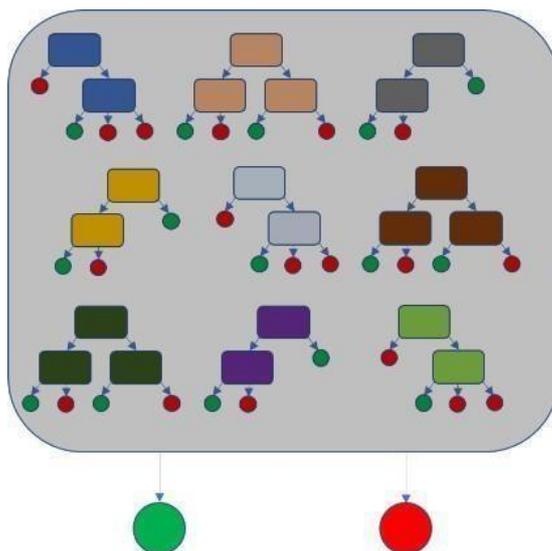
Beberapa macam validasi pada media mainstream yang dilakukan pada penelitian ini yaitu:

- Berita yang disampaikan secara akurat didapatkan dari website [www.kominfo.go.id](http://www.kominfo.go.id).
- Data berita yang sudah dikumpulkan tersebut kemudian akan disimpan dalam format CSV dan akan melewati tahap text preprocessing terlebih dahulu. Di dalam tahap text preprocessing, proses yang pertama kali akan dilakukan yaitu tokenizing. Didalam proses tokenizing ini, semua kalimat dari data yang telah dikumpulkan dipisah berdasarkan tanda spasi.
- Proses folding case. Didalam proses ini, semua huruf besar akan diganti menjadi huruf kecil, dan kata-kata yang mempunyai arti yang sama, akan dinormalisasi menjadi satu kata yang sama.
- Proses filtering. Didalam proses ini, bila terdapat karakter lain kecuali huruf dan angka, maka

- karakter tersebut akan dihapuskan, misalnya karakter tanda baca titik (.).
- e. Proses stopwords removing. Di dalam proses ini, jika terdapat kata-kata yang tidak penting dan tidak unik, maka kata tersebut akan dihapuskan.
  - f. Proses stemming. Di dalam proses ini, semua awalan dan akhiran yang terdapat pada sebuah kata akan dihilangkan, sehingga yang tersisa hanya kata dasarnya saja.
  - g. Tahap terakhir yaitu menghitung bobot pada tiap kata dengan menggunakan pembobotan TF-IDF dengan kombinasi fitur unigram dan bigram.

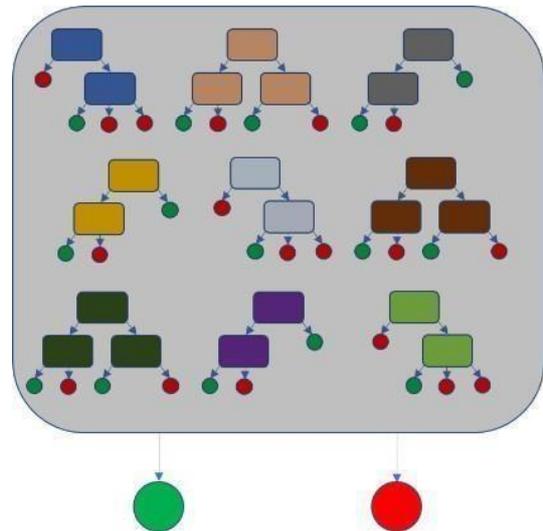
### Klasifikasi Random Forest

Proses klasifikasi yang dilakukan dengan menggunakan metode Random Forest akan digunakan untuk mengelompokkan kategori dari berita hoax (Roihan et al., 2020). Metode Random Forest menerapkan metode bootstrap aggregating (bagging) dan random feature selection (Primajaya & Sari, 2018). Deretan gugus data terdiri dari  $n$  amatan,  $m$  features dan  $p$  peubah penjelas. Berdasarkan 200 data postingan dan komentar yang bersumber dari media sosial Facebook akan dilakukan analisis statistika deskriptif terhadap peubah untuk mendapatkan hasil klasifikasi berita hoax. Secara umum alur kerja Random Forest dapat dilihat pada Gambar 2.



Gambar 2. Alur Kerja Random Forest

Adapun alur kerja implementasi metode Random Forest dalam mendeteksi berita hoax dapat dilihat pada diagram Gambar 3.



Gambar 3. Alur Kerja Implementasi Metode Random Forest

Pada Gambar 3, alur implementasi metode random forest bekerja secara independen untuk membentuk outputnya sendiri dan memberikan hasil prediksi. Metode Random Forest kemudian akan memprediksi dari setiap pohon dan memilih mayoritas kelas yang diprediksi oleh setiap pohon sebagai kelas prediksi sebenarnya dari kumpulan data. Terdapat 3 sampel kalimat berita yang nantinya akan dilakukan proses klasifikasi. Sebelum dilakukan proses klasifikasi, akan dilakukan split dataset terlebih dahulu yang akan menghasilkan vaote pada masing-masing pohon sebelum rata-rata kata yang sesuai dengan berita hoax ditemukan sampai hasil akhir klasifikasi ditemukan.

## HASIL DAN PEMBAHASAN

### Pengumpulan Data Training

Pengumpulan data training dilakukan dengan cara crawling (Putra et al., 2022). Data ini berupa berita yang diambil dari laman website Kominfo dan dikumpulkan menggunakan aplikasi rapid miner serta disimpan kedalam file data dengan format csv. Kemudian data yang telah dikumpulkan tersebut akan digolongkan sentimennya secara manual kedalam dua golongan sentimen yaitu hoax atau fakta pada setiap isi beritanya. Data training yang digunakan dalam penelitian ini yaitu sebanyak 200 data. Data training tersebut dapat dilihat pada Gambar 4.

	A	B
1	comment	VALUE
2	Beredar sebuah unggahan video di media sosial Facebook dengan narasi yang mengklaim bahwa Jepang merayu India agar membantu Australia menyerang Indonesia terkait dengan perebutan Pulau Pa...	hoax
3	Beredar unggahan di media sosial Twitter sebuah video yang menampilkan kondisi jalanan dan bangunan rusak parah pascagempa. Video tersebut diklaim sebagai video gempa yang terjadi di Cianjur, t...	hoax
4	Telah beredar pesan berantai WhatsApp berisi informasi yang mengklaim bahwa Federasi Sepak Bola Dunia (FIFA) membagikan kuota internet gratis sebesar 50 GB untuk menonton pertandingan Piala h...	hoax
5	Beredar di media sosial, sebuah foto yang memperlihatkan deretan bayi di tenda posko gempa Cianjur dengan narasi yang menyebut bahwa bayi tersebut bisa diadopsi karena ibunya menjadi korban n...	hoax
6	Beredar sebuah video di media sosial YouTube dengan klaim Presiden Republik Indonesia Joko Widodo (Jokowi) mengusir Presiden Amerika Serikat Joe Biden yang mengaku nyaman di Bali dan tidak ir...	hoax
7	Beredar di media sosial Facebook, sebuah unggahan video menampilkan gambar mikroskopis dan diklaim sebagai mikrociip robot yang terdapat dalam vaksin Covid-19. Alat itu disebut sebagai nanotekr...	hoax
8	Beredar sebuah akun WhatsApp yang mengatasnamakan Anggota DPRD Kota Sungai Penuh Hutri Randa. Terlihat akun tersebut menghubungi seseorang dengan melakukananrolan di aplikasi WhatsApp ho...	hoax
9	Beredar sebuah video di media sosial Facebook yang membahas keputusan Presiden Joko Widodo atau Jokowi usai penyelenggaraan Konferensi Tingkat Tinggi (KTT) G20 di Bali, pada 15-16 November 2...	hoax
10	Beredar sebuah video di media sosial WhatsApp yang menyebutkan bahwa ada seorang pemuda yang menjadi korban pembegalan di Perempatan Burikan, Kudus.	hoax
11	Beredar sebuah unggahan video rekaman satelit yang memperlihatkan adanya pergerakan tanah yang menyeret rumah dan bangunan lainnya. Video tersebut diklaim sebagai peristiwa yang terjadi saat ho...	hoax
12	Telah beredar pesan berantai WhatsApp sebuah voice note yang mengatakan bahwa akan ada pergeseran lempeng Waduk Cirata di Purwakarta. Informasi tersebut diklaim bersumber dari Badan Metec ho...	hoax
13	Sebuah akun Twitter mengunggah cuitan disertai tautan berita yang menyatakan bahwa Menteri Luar Negeri (Menlu) Rusia Sergei Lavrov setibanya di Bali, dilarikan ke Rumah Sakit (RS) Sanglah, Bali. Pe...	hoax
14	Beredar sebuah unggahan di media sosial Facebook berupa foto lanskap gunung dan mengklaim bahwa gempa yang terjadi di Cianjur disebabkan oleh Gunung Gede dan Gunung Pangrango.	hoax
15	Beredar sebuah informasi yang diklaim berasal dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) bahwa akan terjadi gempa besar di sejumlah wilayah di Sukabumi, Jawa Barat, tepatnya di gar...	hoax
16	Beredar sebuah unggahan di media sosial Facebook yang mengklaim Gunung Gede dan Gunung Pangrango yang berada di Kabupaten Cianjur dan Kabupaten Sukabumi meletus dan mengeluarkan sembo ho...	hoax
17	Beredar di media sosial Facebook, sebuah video berdurasi 8 menit 15 detik yang berisi klaim bahwa Letnan Jenderal Australia memimpin penyerangan ke Indonesia dan dispons dengan keras oleh C...	hoax
18	Beredar di media sosial TikTok, sebuah unggahan video yang memperlihatkan air laut surut dan kering. Unggahan tersebut bertuliskan "ada apa ini seram hari ini pantai Pandangan air laut kering? ser ho...	hoax
19	Beredar sebuah akun WhatsApp yang mengatasnamakan salah satu Anggota DPRD Kabupaten Buleleng Ketut Nguh Arya. Akun tersebut menggunakan nama dan foto, serta terlihat mengirim pesan ke ho...	hoax
20	Beredar sebuah unggahan video yang menampilkan sosok Presiden Rusia Vladimir Putin, Presiden Amerika Serikat Joe Biden, Perdana Menteri Inggris Boris Johnson, dan armada angkatan perang seju...	hoax
21	Telah beredar tangkapan layar sebuah surat yang mengatasnamakan Otoritas Jasa Keuangan (OJK). Surat dengan logo OJK tersebut berisi perihal tanda bukti terdaftar sebagai penyelenggara layanan pe ho...	hoax
22	Beredar sebuah surat pengumuman mengatasnamakan Bank Aceh. Surat tersebut berisi pengumuman kenaikan tarif transfer antarbank dari Rp5000 menjadi Rp150.000 dan pemotongan tarif tersebut ho...	hoax

Gambar 4. Dataset

### Preprocessing Data

Preprocessing data adalah tahapan pembersihan sebelum data diolah lebih lanjut. Tahapan preprocessing data pada penelitian ini menggunakan aplikasi phyton karena penggunaannya yang praktis dan efisien serta dapat dengan mudah diimplementasikan.

Dari data yang telah dikumpulkan sebelumnya, akan dilakukan proses text preprocessing yang dilakukan dalam beberapa tahapan yaitu cleansing, case folding, tokenizing, filtering, dan stemming. Hasil dari tahapan text preprocessing dengan menggunakan phyton dapat dilihat pada Gambar 5.

	after_stemming	VALUE
1	unggah video media sosial facebook narasi klaim jepang rayu india bantu australia serang indonesia kait rebut pulau pasir	hoax
2	unggah media sosial twitter video tampil kondisi jalan bangun rusak parah pascagempa video klaim video gempa cianjur senin 21 november 2022	hoax
3	pesan berantai whatsapp isi informasi klaim federasi sepak bola dunia fifa bagi kuota internet gratis 50 gb tonton tanding piala dunia qatar	hoax
4	media sosial foto deret bayi tenda posko gempa cianjur narasi sebut bayi adopsi ibu korban tinggal dunia akibat gempa cianjur minat rumah sakit daerah rsud sayang cianjur	hoax
5	video media sosial youtube klaim presiden republik indonesia joko widodo jokowi usir presiden amerika serikat joe Biden aku nyaman bal amerika serikat ikut ktt g20	hoax
6	media sosial facebook unggah video tampil gambar mikroskopis klaim mikrociip robot vaksin covid-19 alat nanoteknologi respons sinyal 5g kumpul kirim data	hoax
7	akun whatsapp mengatasnamakan anggota dprd kota sungai penuh hutri randa akun hubung melakukananrolan aplikasi whatsapp	hoax
8	video media sosial facebook bahas putus presiden joko widodo jokowi selenggara konferensi tingkat ktt g20 bal 15-16 november 2022 klaim serta menteri sebut jokowi gagal ktt g20 pecat hormat video s...	hoax
9	video media sosial whatsapp pemuda korban begal empat burik kudus	hoax
10	unggah video rekam satelit gera tanah seret rumah bangun video klaim peristiwa gempa bumi cianjur jawa barat tanggal 21 november 2022	hoax
11	pesan berantai whatsapp voice note geser lempeng waduk cirata purwakarta informasi klaim sumber badan meteorologi klimatologi geofisika bmkg	hoax
12	akun twitter unggah cuit serta taut berita menteri negeri menlu rusia sergei lavrov bal lari rumah sakit rs sang bal jabat lavrov rawat sakit jantung	hoax
13	unggah media sosial facebook foto lanskap gunung klaim gempa cianjur sebab gunung gede gunung pangrango	hoax
14	informasi klaim asal badan meteorologi klimatologi geofisika bmkg gempa wilayah sukabumi jawa barat tepat garis gera sesar cemandiri minggu narasi daerah waspada labuh ratu cibadak cicantayan cib...	hoax
15	unggah media sosial facebook klaim gunung gede gunung pangrango kabupaten cianjur kabupaten sukabumi letus keluar sembur api unggah warga kabupaten cianjur jawa barat kena musibah gempa bu...	hoax
16	media sosial facebook video durasi 8 menit 15 detik isi klaim letnan jenderal australia pimpin serang indonesia respons keras china video tampil letnan jenderal australia richard maxwell burr serta narasi j...	hoax
17	media sosial tiktok unggah video air laut surut kering unggah tulis seram pantai pandangan air laut kering waspada	hoax
18	akun whatsapp mengatasnamakan salah anggota dprd kabupaten buleleng ketut nguh arya akun nama foto kirim pesan kerabat modus tipu	hoax
19	unggah video tampil sosok presiden rusia vladimir putin presiden amerika serikat joe Biden perdana menteri inggris bor johnson armada angkat perang negara video klaim bentuk dukung presiden rusia v...	hoax

Gambar 5. Hasil Tahapan Text Processing Dengan Menggunakan Phyton

### Pengujian Dengan Data Uji

Tahapan pengujian merupakan tahapan yang utama didalam proses analisis data. Pada penelitian ini, tahapan pengujian dilakukan dengan menggunakan bahasa pemrograman Python yang dijalankan melalui Jupyter Notebook. Data training yang sebelumnya telah dibersihkan dan diberikan bobotnya akan diunggah terlebih dahulu dan diimport datanya kedalam program menggunakan library pandas. Kemudian data training tersebut dibagi kedalam dua variabel yaitu hoax dan fakta.

Variabel hoax merupakan variabel yang menampung bobot dari tiap kata dan variabel fakta merupakan variabel yang menampung hasil sentimen dari tiap berita yang ada. Kemudian data uji diinputkan

kedalam program yang sudah tersedia. Data yang diinputkan merupakan data dari hasil pembobotan. Kemudian, program akan menggolongkan berita menggunakan metode Random Forest dengan memanfaatkan library sklearn yang tersedia pada Python. Output yang dihasilkan merupakan nilai akurasi dari proses pengujian dengan menggunakan data uji. Dalam penelitian ini, data yang digunakan sebagai data training ada sebanyak 100 data dan data yang digunakan sebagai data uji ada sebanyak 100 data. Hasil pengujian dari data uji tersebut dapat dilihat pada Gambar 6.

	content	aktual	prediksi			
1						
2	akun face	hoax	hoax	1		
3	media sos	hoax	hoax	1		
4	video me	hoax	hoax	1		
5	lansir lipu	fakta	fakta	1		
6	lansir kor	fakta	fakta	1		
7	lansir kon	fakta	fakta	1		
8	media sos	hoax	hoax	1		
9	lansir mei	fakta	fakta	1		
10	akun wha	fakta	fakta	1		
11	media sos	hoax	hoax	1		
12	video me	hoax	hoax	1		
13	lansir kab	fakta	fakta	1		
14	akun wha	fakta	fakta	1		
15	media sos	hoax	hoax	1		
16	sinar mas	fakta	fakta	1		
17	akun face	fakta	fakta	1		
18	pamflet n	hoax	hoax	1		
19	lansir me	fakta	fakta	1		
20	pesan be	hoax	hoax	1		
21	akun wha	hoax	hoax	1		
22	klaim das	fakta	fakta	1		
23	media sos	hoax	hoax	1		
24	klaim vid	fakta	fakta	1		
25	unggah vi	hoax	hoax	1		
26	akun wha	hoax	hoax	1		

Gambar 6. Hasil Pengujian Data Uji

**Evaluation Measure**

Evaluation Measure merupakan tahapan yang dilakukan untuk menguji hasil klasifikasi dari sistem dengan cara menguji nilai kebenarannya (Khan et al., 2021). Data set dibagi kedalam dua kelompok kelas yaitu hoax dan fakta. Data yang digunakan untuk pengukuran evaluasi pada penelitian ini adalah sebanyak 200 data. Data tersebut kemudian akan diuji dengan menggunakan evaluation measure sehingga diperoleh nilai akurasi. Hasil dari perhitungan dengan menggunakan evaluation measure dapat dilihat pada Gambar 7.

	precision	recall	f1-score	support
hoax	1.00	1.00	1.00	21
accuracy			1.00	21
macro avg	1.00	1.00	1.00	21
weighted avg	1.00	1.00	1.00	21
confusion matrix: [[21]]				
-----				
	precision	recall	f1-score	support
hoax	1.00	1.00	1.00	20
accuracy			1.00	20
macro avg	1.00	1.00	1.00	20
weighted avg	1.00	1.00	1.00	20
confusion matrix: [[20]]				
-----				

Gambar 7. Evaluation Measure

Dari hasil pengujian sistem diperoleh nilai akurasi sebesar 0.9803 atau 98%.

**KESIMPULAN**

Berdasarkan hasil penelitian yang sudah dilakukan, maka kesimpulan yang dapat ditarik yaitu metode Random Forest terbukti dapat menggolongkan sentimen berita berdasarkan data yang telah didapatkan. Uji coba yang dilakukan dengan menggunakan data testing didapatkan hasil berupa sentiment positif sebanyak 28 dan sentimen negatif sebanyak 23. Dari perhitungan menggunakan evaluation measure, dimana nilai k optimalnya = 3 dan data yang digunakan sebanyak 80% data training dan 20% data testing dari 200 dataset diperoleh hasil accuracy sebesar 98%, precision sebesar 50%, recall sebesar 100% dan F1-Score sebesar 89%.

**DAFTAR PUSTAKA**

Bahri, S. (2021). Literasi Digital Menangkal Hoaks Covid-19 Di Media Sosial. *Jurnal Ilmu Komunikasi*, 10(1), 16–28.

Edy, S., & Florina, I. D. (2021). ANALISIS SEBARAN HOAKS PADA FACEBOOK PERIHAL INFO SEPUTAR COVID-19. *JURNAL SIGNAL*, 9(2), 344. <https://doi.org/10.33603/signal.v9i2.6290>

Khan, M. Y., Qayoom, A., Nizami, M. S., Siddiqui, M. S., Wasi, S., & Raazi, S. M. K.-R. (2021). Automated Prediction of Good Dictionary EXamples (GDEX): A Comprehensive Experiment with Distant Supervision, Machine Learning, and Word Embedding-Based Deep Learning Techniques. *Complexity*, 2021, 1–18. <https://doi.org/10.1155/2021/2553199>

Manalu, D. R., L. Tobing, M. C., & Yohanna, M. (2022). Analisis Sentimen Twitter Terhadap Wacana Penundaan Pemilu Dengan Metode Support Vector Machine. *METHOMIKA Jurnal Manajemen Informatika Dan Komputerisasi Akuntansi*, 6(6), 149–156. <https://doi.org/10.46880/jmika.Vol6No2.pp149-156>

Primajaya, A., & Sari, B. N. (2018). Random Forest Algorithm for Prediction of Precipitation. *Indonesian Journal of Artificial Intelligence and Data Mining*, 1(1), 27. <https://doi.org/10.24014/ijaidm.v1i1.4903>

Putra, P., Indra, I. B., & Eniyati, S. (2022). Analisis Pola Pembelian Konsumen pada Data Transaksi Penjualan Suku Cadang Mobil dengan Algoritma FP-Growth (Studi Kasus: PT. Sun Star Motor Kudus). *Jurnal Ilmiah Universitas Batanghari Jambi*, 22(2), 882. <https://doi.org/10.33087/jiubj.v22i2.2004>

Roflin, E., Andriyanu, I., & Liberti, P. (2021). Penaksiran rataan dan variasi populasi pada sampel acak terstratifikasi dengan auxiliary variabele. *Intelektiva : Jurnal Ekonomi, Sosial & Humaniora*, 2(7).

Roihan, A., Sunarya, P. A., & Rafika, A. S. (2020).

Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper. *IJCIT (Indonesian Journal on Computer and Information Technology)*, 5(1).

<https://doi.org/10.31294/ijcit.v5i1.7951>

Susanti, N. A., Walid, M., & Hoiriyah, H. (2022).

Klasifikasi Data Tweet Ujaran Kebencian di Media Sosial Menggunakan Naive Bayes Classifier. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(2), 538–543.

<https://doi.org/10.36040/jati.v6i2.5174>